

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2000, 17:37:21 ; Search time 23795.2 Seconds  
(without alignments)  
973.543 Million cell updates/sec

Title: US-09-227-881-2

Sequence: 1 atcttgcgttcgttaccctc.....caggcacctctcagcacagc 5304

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1033670 seqs, 218378903 residues

Total number of hits satisfying chosen parameters: 2067340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_om: \*  
4: gb\_ov: \*  
5: gb\_pac: \*  
6: gb\_ph: \*  
7: gb\_pl1: \*  
8: gb\_pl2: \*  
9: gb\_pr1: \*  
10: gb\_pr2: \*  
11: gb\_pr3: \*  
12: gb\_ro: \*  
13: gb\_sy: \*  
14: gb\_un: \*  
15: em\_fun: \*  
16: em\_hum1: \*  
17: em\_hum2: \*  
18: em\_in: \*  
19: em\_om: \*  
20: em\_or: \*  
21: em\_ov: \*  
22: em\_pat: \*  
23: em\_ph: \*  
24: em\_pl: \*  
25: em\_ro: \*  
26: em\_sts: \*  
27: em\_sy: \*  
28: em\_un: \*  
29: em\_v1: \*  
30: gb\_ba3: \*  
31: gb\_in1: \*  
32: gb\_in2: \*  
33: gb\_in3: \*  
34: gb\_pl3: \*  
35: gb\_pr4: \*  
36: em\_ba1: \*  
37: em\_ba2: \*  
38: em\_hc1: \*  
39: em\_hc2: \*  
40: em\_hc3: \*  
41: em\_hc4: \*  
42: em\_hc5: \*  
43: em\_hc6: \*

44: em\_hc7: \*  
45: em\_hc8: \*  
46: em\_hc9: \*  
47: em\_hc10: \*  
48: em\_hc11: \*  
49: em\_hc12: \*  
50: em\_hc13: \*  
51: em\_hc14: \*  
52: em\_hc15: \*  
53: em\_hc16: \*  
54: em\_hc17: \*  
55: em\_hc18: \*  
56: em\_hc19: \*  
57: em\_hc20: \*  
58: em\_hc21: \*  
59: em\_hc22: \*  
60: em\_hc23: \*  
61: em\_hum3: \*  
62: em\_hum4: \*  
63: em\_hum5: \*  
64: em\_hum6: \*  
65: gb\_pr5: \*  
66: gb\_pr6: \*  
67: gb\_pr7: \*  
68: gb\_hc1: \*  
69: gb\_hc2: \*  
70: gb\_hc3: \*  
71: gb\_hc4: \*  
72: gb\_hc5: \*  
73: gb\_hc6: \*  
74: gb\_hc7: \*  
75: gb\_hc8: \*  
76: gb\_hc9: \*  
77: gb\_hc10: \*  
78: gb\_hc11: \*  
79: gb\_hc12: \*  
80: gb\_hc13: \*  
81: gb\_hc14: \*  
82: gb\_hc15: \*  
83: gb\_hc16: \*  
84: gb\_hc17: \*  
85: gb\_hc18: \*  
86: gb\_hc19: \*  
87: gb\_hc20: \*  
88: gb\_hc21: \*  
89: gb\_hc22: \*  
90: gb\_hc23: \*  
91: gb\_sts1: \*  
92: gb\_sts2: \*  
93: gb\_v11: \*  
94: gb\_v12: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	5274	99.4	79376	65 HS454G6	298750 Human DNA s
2	5253.4	99.0	5300	11 AF007562	AF007562 Homo sapi
3	5223	98.5	170425	77 AC024490	AC024490 Homo sapi
4	1836	34.6	2800	66 HSMYOC1	AF049791 Homo sapi
5	417.8	7.9	1228	9 AB006686S1	AB006686 Homo sapi
6	407.8	7.7	1086	66 HSGLC1A1	297171 Homo sapien
7	189.8	3.6	161577	10 AC007688	AC007688 Homo sapi
8	189.8	3.6	193123	77 AC023790	AC023790 Homo sapi
9	189	3.6	199722	71 AC012404	AC012404 Homo sapi
C 10	187.8	3.5	97037	9 AC004973	AC004973 Homo sapi
C 11	187.8	3.5	135038	67 HUMYXWD703	L78810 Homo sapien
C 12	185.8	3.5	76727	65 HS821D11	AL021453 Human DNA

C	13	185.8	3.5	152044	79	AC026395	AC026395 Homo sapi
C	14	185.8	3.5	157057	72	AC025947	AC025947 Homo sapi
C	15	185.8	3.5	161499	78	AC015488	AC015488 Homo sapi
C	16	185.8	3.5	184656	86	AL161726	AL161726 Homo sapi
C	17	185.8	3.5	200681	69	AC008755	AC008755 Homo sapi
C	18	185.6	3.5	201372	86	AL157941	AL157941 Homo sapi
C	19	184.8	3.5	146190	89	AP001813	AP001813 Homo sapi
C	20	184.8	3.5	163494	89	AP002391	AP002391 Homo sapi
C	21	184.8	3.5	183241	89	AP001354	AP001354 Homo sapi
C	22	184.6	3.5	176029	11	AC011362	AC011362 Homo sapi
C	23	184.2	3.5	130020	67	HUAC004525	AC004525 Homo sapi
C	24	184	3.5	157304	73	AC024720	AC024720 Homo sapi
C	25	184	3.5	187709	73	AC016168	AC016168 Homo sapi
C	26	184	3.5	233734	70	AC011407	AC011407 Homo sapi
C	27	183.6	3.5	62070	88	AL358852	AL358852 Homo sapi
C	28	183.6	3.5	129370	89	AP000609	AP000609 Homo sapi
C	29	183.6	3.5	169333	89	AP002789	AP002789 Homo sapi
C	30	183.6	3.5	193171	74	AC018723	AC018723 Homo sapi
C	31	183.4	3.5	41407	11	AC053467	AC053467 Homo sapi
C	32	183.4	3.5	149138	79	AC026936	AC026936 Homo sapi
C	33	183.4	3.5	156331	90	HS193M1	AL121715 Homo sapi
C	34	183.4	3.5	174974	87	AL356055	AL356055 Homo sapi
C	35	183.4	3.5	195364	65	HS431A14	283956 Human DNA S
C	36	183.4	3.5	198583	74	AC019114	AC019114 Homo sapi
C	37	183.2	3.5	41936	35	CH19R30879	AD000684 Homo sapi
C	38	183.2	3.5	19128	10	AC007536	AC007536 Homo sapi
C	39	183.2	3.5	119483	9	AC005588	AC005588 Homo sapi
C	40	183.2	3.5	145528	9	AC003665	AC003665 Homo sapi
C	41	183.2	3.5	167943	65	HS267M20	AL031053 Human DNA
C	42	183.2	3.5	170245	65	HS109F14	AL022721 Human DNA
C	43	183	3.5	42686	9	AC000093	AC000093 Homo sapi
C	44	183	3.5	91767	10	AC007748	AC007748 Homo sapi
C	45	183	3.5	112748	10	AC007242	AC007242 Homo sapi

## ALIGNMENTS

RESULT 1  
 HS4546/c  
 LOCUS  
 DEFINITION Human DNA sequence from PAC 45466 on chromosome 1q24. Contains  
 trabecular meshwork inducible glucocorticoid response protein,  
 TIGR, myocilin, ESTs and STS.  
 ACCESSION 298750  
 VERSION 1q24: myocilin; TIGR.  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 79376)  
 AUTHORS Deadman,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-OCT-1997) Chromosome 1 Project Group  
 (http://www.sanger.ac.uk/HGP/Chr1/) Sanger Centre, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humguery@sanger.ac.uk Clone requests: clones@esanger.ac.uk  
 On Feb 14, 1998 this sequence version replaced gi:245500.  
 IMPORTANT: This sequence is not the entire insert of clone 45466.  
 It may be shorter because we only sequence overlapping sections  
 once, or longer because we arrange for a small overlap between  
 neighbouring submissions.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variations annotated may not be found in the sequence submission  
 corresponding to the overlapping clone as we submit sequences with  
 only a small overlap as described above.  
 This sequence was generated from part of bacterial clone contigs of  
 human chromosome 1, constructed by the Sanger Centre chromosome 1  
 mapping group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr1/

This sequence has been finished according to sequence map criteria  
 as follows. An attempt is made to resolve all sequencing problems,  
 such as compressions and repeats, but not necessarily within known  
 annotated human repeat sequence elements (e.g. Alu). Where the  
 sequence is ambiguous, there is an annotation using the 'unsure'  
 feature key.  
 The true left end of clone 45466 is at 1 in this sequence. The true  
 left end of clone 56089 is at 79273.  
 45466 is from the library RPEC3 constructed at the Roswell Park  
 Cancer Institute by the group of Pieter de Jong.  
 For further details see <http://pacpac.med.buffalo.edu/>.

## FEATURES

source  
 1..79376  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /map="1q24"  
 /clone="RP3-45466"  
 /clone\_1lb="RPECI-3"  
 repeat\_region  
 435..472  
 /note="19 copies of 2 mer 82 & conserved"  
 prim\_transcript  
 1914..>3968  
 /note="match: multiple ESTs  
 match: R56676 AA043968 W63639 F12081 AA046699  
 match: F02925 AA131540 W00634 R36066 AA13383  
 match: AA163561 F02925 AA131540 W00634 R36066  
 match: AA13383 AA163561 N891740 AA17481 AA057059  
 match: AA439084 W47082 AA043965 AA341783 AA353681  
 match: AA046487 AA369741 H08313 AA166899 H327230  
 match: H08333 H08236 N42052 D61944 R27102 N32353  
 match: N30491 AA307150 AA192"  
 3703..3746  
 /note="22 copies of 2 mer 89 & conserved"  
 repeat\_region  
 4051..4183  
 /note="AluSq repeat: matches 1..133 of consensus  
 incomplete repeat"  
 repeat\_region  
 4200..4502  
 /note="AluSq repeat: matches 2..301 of consensus"  
 4659..4851  
 /note="AluSx repeat: matches 2..194 of consensus  
 incomplete repeat"  
 repeat\_region  
 5216..5345  
 /note="AluJo repeat: matches 132..1 of consensus  
 incomplete repeat"  
 repeat\_region  
 7759..7907  
 /note="MIR repeat: matches 174..1 of consensus"  
 7933..9328  
 /note="TIGER1 repeat: matches 1..1472 of consensus"  
 9332..9626  
 /note="AluSg repeat: matches 1..289 of consensus"  
 9639..10335  
 /note="TIGER1 repeat: matches 1469..2174 of consensus"  
 10343..10642  
 /note="AluSP repeat: matches 1..300 of consensus"  
 10643..10836  
 /note="TIGER1 repeat: matches 2175..2417 of consensus"  
 <10903..>16855  
 /note="endogenous retroviral sequence"  
 10949..11384  
 /note="LTR2 repeat: matches 31..449 of consensus"  
 15835..15870  
 /note="18 copies of 2 mer 83 & conserved"  
 16856..17286  
 /note="LTR2 repeat: matches 31..449 of consensus"  
 17287..17574  
 /note="AluSg repeat: matches 15..300 of consensus"  
 18294..18650  
 /note="THER1B repeat: matches 358..1 of consensus"  
 18877..19180  
 /note="AluY repeat: matches 301..2 of consensus"  
 19767..20013  
 /note="LTR2 internal repeat: matches 5002..4750 of  
 consensus"

repeat_region	20051..20118	/note="MUT1F repeat: matches 539..471 of consensus"
repeat_region	20130..20261	/note="Aluub repeat: matches 131..1 of consensus incomplete repeat"
repeat_region	20264..20722	/note="MUT1F repeat: matches 482..1 of consensus"
repeat_region	20858..21223	/note="MST2 repeat: matches 394..1 of consensus"
repeat_region	21216..21302	/note="MUT2_internal repeat: matches 4520..443 of consensus"
repeat_region	21403..21703	/note="MUT2_internal repeat: matches 3887..3580 of consensus"
repeat_region	21978..22357	/note="MUT2_internal repeat: matches 3218..2839 of consensus"
repeat_region	22363..22524	/note="MUT2_internal repeat: matches 2495..2317 of consensus"
repeat_region	22531..22839	/note="Alusx repeat: matches 302..1 of consensus"
repeat_region	23007..23309	/note="Alusp repeat: matches 3..301 of consensus"
misc_feature	23286..>23580	/note="STS G07544"
misc_feature	23580..24072	/note="STS G07436"
repeat_region	24008..24280	/note="Aluub repeat: matches 298..6 of consensus"
repeat_region	23895..26364	/note="LTR2 repeat: matches 2..449 of consensus"
repeat_region	26397..26697	/note="Alusg repeat: matches 1..302 of consensus"
repeat_region	26719..27021	/note="Alusx repeat: matches 1..302 of consensus"
repeat_region	29037..29334	/note="Alusx repeat: matches 300..3 of consensus"
repeat_region	30028..30310	/note="Alusx repeat: matches 1..301 of consensus"
repeat_region	30402..30559	/note="MUT2B repeat: matches 264..404 of consensus"
repeat_region	30402..30615	/note="MUT2A repeat: matches 264..453 of consensus"
prim_transcript	>30801..>31136	/note="match: 5' EST AA256902 clone 682136"
repeat_region	30980..31222	/note="MWR4B repeat: matches 199..431 of consensus"
repeat_region	31232..31528	/note="Alusc repeat: matches 299..1 of consensus"
repeat_region	31530..31654	/note="MWR4A repeat: matches 417..540 of consensus"
repeat_region	31836..32135	/note="Aluy repeat: matches 300..1 of consensus"
repeat_region	32200..32301	/note="Alufo repeat: matches 186..287 of consensus incomplete repeat"
repeat_region	32365..32493	/note="MWR repeat: matches 206..67 of consensus"
repeat_region	33283..33399	/note="MWR repeat: matches 77..196 of consensus"
repeat_region	33408..33667	/note="Aluy repeat: matches 1..291 of consensus"
repeat_region	33956..34043	/note="MWR repeat: matches 80..167 of consensus"
repeat_region	34725..34904	/note="Aluob repeat: matches 299..127 of consensus incomplete repeat"
repeat_region	34907..35207	/note="Aluy repeat: matches 300..1 of consensus"
repeat_region	35212..35344	/note="Aluub repeat: matches 133..1 of consensus"

[illegible]

Oy	721	atgcacgtggctaaagccctggaccttccaagggaataatgaataacccgagagcaaaacaa	780
Db	67686	ATTGACTGGGCTTAACCCCTGGACCTTTCAAGGAAATATGAAAAACGTAGGCAAAAACAAA	676227
Oy	781	gacacgtgttaaaaggcaaacacagaacatgtgagccctccaagcagcagtgccccctcagca	840
Db	67626	GACATGTGTTAAAGGCAACCAACATGTGTAGCCTTCAAAACACAGAGGCCCTCAGCA	67567
Oy	841	gggacccctgaggcaattgcctttaaggagccagtttctcttaaggaatcttaagaaatc	900
Db	67566	GGGACCCCTGAGGCATTGTGCTTTAGGAAGGCCAGTTTCTTAAGGAATCTTAAGAAACAC	67507
Oy	901	tggaaagctcaagaaatttaaacattttaagtataaacaatatgacgaatgatatatcag	960
Db	67446	TTTAAACATAGGTGCTCCCAATTTTATTAAGCAGGACATACAAAGATTAACGTCGCCAGCTCC	67387
Oy	961	tttaagacatgggtgcccaatttataaagtcaggaatacaaggaataacgtgtgccagctcc	1020
Db	67446	TTTAAACATAGGTGCTCCCAATTTTATTAAGCAGGACATACAAAGATTAACGTCGCCAGCTCC	67387
Oy	1021	ggataggttcaggaatacattagaaatcacctgtgtcccatccctaacttttccgaatgatac	1080
Db	67386	GGATAGGTTCAGGAATCATTTAGAAATCACGTGTGCCCATCTTAACCTTTTCAGAAATGATC	67327
Oy	1081	tgtcaatgccccacacaacaagcccgatgtgtctgacatacaacacatctacaaccga	1140
Db	67326	TGTCAATAGCCCTCACACACAGGCCCGGATGTGTCTAGACCTACAAACCAATCTAACACCGAA	67267
Oy	1141	gtgctcaacacatgtgttaacgtgtcaatctagtaagttcccatlacaatatgcacctccc	1200
Db	67286	GTGCTCAACCATTTGTTAAACGTGTATCTCAATAGTCCCATTAACAAAGCCACCTCCCC	67207
Oy	1201	tgtgagcccatcccgctccacaggaagctcccaactctagaactctgcatacagaatgt	1260
Db	67206	TGTGAGCCCATCCCGCTCCACAGGAAGTCTCCCACTGTAACTTGTGCATCAGATGT	67147
Oy	1261	tacagccgaagctcccgtagagggtgaggtctgtgtcttaacaactctgatatgtcttac	1320
Db	67146	TACACCCAGAACCTCGTAGGGGTGAGGGTCTGTGTCTTACACCTAATCGTATGCTCTAC	67087
Oy	1321	acctgaagctcaatgtgaaacctgcgctcccaaggttcaagaaatctccctgtccaaagctcc	1380
Db	67086	ACCTAAGCTCACTGCNAACCTCTGCTCCAGGTTCAAGCAATTTCTCTGTCTCAGCCTCC	67027
Oy	1381	cgcgtagctctgagacatacagcgacgaccccggttaatttttgatgttgtagaagatggg	1440
Db	67026	CGCGTAGCTGAGACTACAGGCGCACGCCGCGCTAATTTTGTGATTTGTAGTAGAGATGGG	66967
Oy	1441	gtttcaacaatatagcccggtctgtgtcttgaactctctgacctgaagttgatccaaccatc	1500
Db	66966	GTTTACCATATATTACCCGGCGTGTGTAACCTCTGACCTAGGTGAATCAACCAACCTC	66907
Oy	1501	agccctccaaagtgcttgagattatcaaggacatgagtaacgcgcccgccgaagggtcaatgt	1560
Db	66906	AGCCCTCTTAAAGCTGTGGATTACAGGCATGAGTCAACGCCGCCGACCAAGGCTAGTGT	66847
Oy	1561	tttaataaggaataactctgaatggttttactaaaccaaacaggggaacagaacaaagctgtga	1620
Db	66846	TTAATAAGGAATAACTTGATGATGTTTACTTAACCAACAAAGGAACAGACAAAGCTGTGA	66787
Oy	1621	ttaattcagggttctcttggaatgggggaatgtgtgcatagtgcgtgcctagttcccaagc	1680
Db	66786	TAAATTTCAGGGATTTCTTGGGATGGGGAATGGGCATGAGCTGCTGCTTAATCCAGAC	66727
Oy	1681	caatgtccctcaacacttctctccctcatccctcaatttccaagctaaatlatcatcttatt	1740
Db	66726	CACGTGCTCTACTACTTTCTTCCCTCATCCCATTTTTCAGGCTAAGTTACATTTTAT	66667
Oy	1741	cacacatgctcttgvtgtaagcctccacatcgttactgaaataagatatataaactag	1800
Db	66666	CACCATGCTTTGTGTGTAGCCTCCACATCGTTACTGAAATAGATGTATACCTTAACATAG	66607

QY	1801	ttccatttgggggcacatctgtgtgtgtgtatagagggaaggaagcacaaccagaagactcct	1860
Db	66606	tttccatttgggggcacatctgtgtgtgtgtatagagggaagggcacaatcccaagactcct	66547
QY	1861	tgaagccccccagagaggttctctctccagctggggggagccccctgtcaagcaccggagtcc	1920
Db	66546	tgaaagccccccagagaggttctctctccagctggggggagccccctgtcaagcaccggagtcc	66487
QY	1921	tgggtgtccctgagcaaacctgtccagccgtgtccacttgttltgttlatcactctcag	1980
Db	66486	tgggtgtccctgagcaaacctgtccagccgtgtccacttgttltgttlatcactctcag	66427
QY	1981	gaacctgtgtcttccattctgtgtgtactgttcatctcaaggacatctcatgtgaact	2040
Db	66426	gaacctgtgtcttccattctgtgtgtactgttcatctcaaggacatctcatgtgaact	66367
QY	2041	tatttgatctatatctgcgcagacacccagagacaaaatgggtgagcaagagctacatgc	2100
Db	66366	tatttgatctatatctgcgcagacacccagagacaaaatgggtgagcaagagctacatgc	66307
QY	2101	ccctaccctctgtgaggtgacaagtcttctcaatggaagacgtgcagaagaaataatagcca	2160
Db	66306	ccttaccttctgtgaggtgacaagtcttctcaatggaagacgtgcagaagaaataatagcca	66247
QY	2161	gccaacttaaacccasgtgctggaagaagaataacaacatcttgaagaattgtgcg	2220
Db	66246	gccaaacttmaacccasgtgctggaagaagaataaacacatcttgaagaattgtgcgc	66187
QY	2221	agcatcccttaacaagagccacactccctcagagcccccctgtctcatcatgtgtccggag	2280
Db	66186	agcatcccttmaacagagccacactccctcagagcccccctgtctcatcatgtgtccggag	66127
QY	2281	cccccaagcccgagcttctccaaagcctctcctccatacagtcacagcgtctcagcttgcct	2340
Db	66126	cccccaagcccgagcttctccaaagcctctcctccatacagtcacagcgtctcagcttgcct	66067
QY	2341	gctcgtcttcccgatcgttcctgtgtgcatctgagcttggagaactccttgtctcagact	2400
Db	66066	gctcgtcttcccgatcgttcctgtgtgcatctgagcttggagaactccttgtctcagact	66007
QY	2401	ccagaagaagaaatggaaggaggaaactatctctcagagaaatcttggaagggaagcttcc	2460
Db	66006	ccagaagaagaaatggaaggaggaaactatctctcagagaaatcttggaagggaagcttcc	65947
QY	2461	ctcagaagggaaggggccctccacgctccagaagaaatctcgaagggttggggagctcagggg	2520
Db	65946	ctcagaagggaaggggccctccacgctccagaagaaatctcgaagggttggggagctcagggg	65887
QY	2521	tggggagcgtggggctgtgagcgggtgcctgaagcgaggaagtgtgaaggcgaaagctga	2580
Db	65886	tggggagcgtggggctgtgagcgggtgcctgaagcgaggaagtgtgaaggcgaaagctga	65827
QY	2581	gctgcgccagatgttcaagtgtgttcaaggggcttgggaagtttcctgtcttccctgtgag	2640
Db	65826	gctgcgccagatgttcaagtgtgttcaaggggcttgggaagtttcctgtcttccctgtgag	65767
QY	2641	cttttatacttttccctctgtctgtgagaggaagaagcttatcttcaagaagatgcagtttc	2700
Db	65766	cttttatacttttccctctgtgtgtgagaggaagaagcttatcttcaagaagatgcagtttc	65707
QY	2701	ataaagtcagctgttataaatctccagagtggtgatgggttcttccctcacaagagccttat	2760
Db	65706	ataaagtcagctgttataaatctccagagtggtgatgggttcttccctcacaagagccttat	65647
QY	2761	ttaatgggaatacagaagcgaagctcatcttccaaagccgttaattcaaggaagaagtga	2820
Db	65646	ttaatgggaatacagaagcgaagctcatcttccaaagccgttaattcaaggaagaagtga	65587
QY	2821	tgggaatcttttcttcaagtcttcttgggcaactactgaagccctgtgtggagcttgctta	2880
Db	65586	tgggaatcttttcttcaagtcttcttgggcaactactgaagccctgtgtgtggagcttgctta	65527
QY	2881	tgcagaagcgttcgaaaaccttggaaatcaggaacctcggttcttcttctgtgtctgcat	2940

```
|||||
Db 65526 TGCAGACGGTGCAGAAACCTTGGATCAGGAGCTCGTTTCTTCTGTTGTCGCATT 65467
Qy 2941 ggttgctgtgcaacgctgggcaagtgctctctctccctgggccaagtctctctgct 3000
Db 65466 GGTGTGGCTGTGCACCTCGGGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 65407
Qy 3001 ataaagacccttcagctctcgtgtctgtgtgaacacctccctgtgactctctgtgagg 3060
Db 65406 ATAAAGACCTTGCAGCTCTGTTGTGTGTGAACCTTCCCTGTGATTTCTGTGAGGGG 65347
Qy 3061 ggaatgtgaaggaggaaaggagagagctgagagctgagccagagaggaggaggagg 3120
Db 65346 GGAATGTTGAAGGGAGGAGAGGAGGAGAGCTGGAGAGCTGAGCCAGAGGGAGGGAGGG 65287
Qy 3121 ggaacaggaaggagagagagagctgggtgtctccatcagctccctactgtactgaagtc 3180
Db 65286 GGACAGAGAGGAGGAGGAGAGAGCTGGGTGCTCCATCACTCTCTCTCTCTCTCTCTCT 65227
Qy 3181 caggaacaggaagagagagagcttcaggaagagctgaatgaacccagagccagcttccct 3240
Db 65226 CAGGACCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 65167
Qy 3241 tccctaagcaagaacagctgcaatctgccaataaccaaagaatgagagagactaactgtc 3300
Db 65166 TCCCTAAGCATAGACATGGCATTTGCCAATACCAAAAGAAATGAGAGACTAATCGT 65107
Qy 3301 ggtagctcttctgctgcaatcacaacactgggagagagagctggaagaaatggcagaatg 3360
Db 65106 GGTAGCTTTTGTCTGTGCTGCTCAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 65047
Qy 3361 ttaaaccttcacaccttcagagagagagagagagagagagagagagagagagagagag 3420
Db 65046 TTAACCTTTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCT 64987
Qy 3421 agtgaacctgagagagagagagagagagagagagagagagagagagagagagagagag 3480
Db 64986 AGTGAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 64927
Qy 3481 acagatcattcaagagagagagagagagagagagagagagagagagagagagagagagag 3540
Db 64926 ACAGATTCATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 64867
Qy 3541 gttcagaagagagagagagagagagagagagagagagagagagagagagagagagagag 3600
Db 64866 GTTCTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 64807
Qy 3601 cctgaattcctaatactatcttctccttcaagaagctgagtaattctgagcaagctcagaag 3660
Db 64806 CCGATTTCTAATATCTAATATTTTCTTCAAGCTGAGTAATCTGAGCAAGTCAACAG 64747
Qy 3661 gtatgaactgaagctgaagaatcactgaatctcctcctataagaaactcttctctgtc 3720
Db 64746 GTATGAATCTAGGCTGTAGGTTACTTACTTACTTACTTACTTACTTACTTACTTACT 64687
Qy 3721 ggaagttagcagagagagagagagagagagagagagagagagagagagagagagagag 3780
Db 64686 GGAGTTAGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 64627
Qy 3781 taagaacaaacagagctcaagagagagagagagagagagagagagagagagagagagag 3840
Db 64626 TAAAGCAAAAGATTCAGGCTAGGCTTCTGCTGACTAATAGATGGTTTTTGAAGAAAT 64567
Qy 3841 caattcagagagagagagagagagagagagagagagagagagagagagagagagagagag 3900
Db 64566 CATTTGCGGATGTCTTACTAATCTGATTCAGAAATGAGACAGTACCTTTGGTCAGTG 64507
Qy 3901 taagaacaaacagagagagagagagagagagagagagagagagagagagagagagagag 3960
Db 64506 TAAACAAACACCTTTGTAAATGTCTCAAGTTAAGGCTTAAGCTGACAGAACCAATCAAT 64447
Qy 3961 aagagataagactcttagagagagagagagagagagagagagagagagagagagagagag 4020
|||||
|||||
Db 64446 AAGAAATGAATCTTTAGAGCAAACTGTGTTCTCCACTCGAGAGTGTGCTCCAGGGC 64387
Qy 4021 agtttgaagaaatatttacttcaagaagatgacacigtgtgtgtatataaacaataaagt 4080
Db 64386 AGTTTGAAGAAATTTTACTTACAGATGATGACACTGTGTGTGTGTGTGTGTGTGTGT 64327
Qy 4081 tgcctcaagagcaatcatatctcaagtgagctaaagttactcttcagagcttctgtatcat 4140
Db 64326 TGCTCAAGGCATCATTAATTTCAAGGCGTTAAAGTTACTTCTGACAGTTTGGTATAT 64267
Qy 4141 ttaatggctatggccatttgcttctgtcttctcctcttgggttctatcaatgaagcag 4200
Db 64266 TTAATGGCTATTGCGCATTTTCTTTTGTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 64207
Qy 4201 ggaattataacccaagctccagaagagagagagagagagagagagagagagagagagag 4260
Db 64206 GGAATTAATTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCT 64147
Qy 4261 taatttaccaccccttcaactaaatlaacatctlaattcattcagatgagagagagagag 4320
Db 64146 TGTTTTACCACTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 64087
Qy 4321 ccaaaagtgtaataagagagagagagagagagagagagagagagagagagagagagagag 4380
Db 64086 CTCAAAAGTGAATTAACAGTACCTGTGATTTGTGATTTGATTTGATTTGATTTGATTT 64027
Qy 4381 ttatacattatcaagtggtgtgcaagagagagagagagagagagagagagagagagagagag 4440
Db 64026 TTAACTATATTAACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 63967
Qy 4441 ctttgaagatlagagagagagagagagagagagagagagagagagagagagagagagag 4500
Db 63966 CTTTGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 63907
Qy 4501 attttgataattgagaaatcaatattcaatcaatctgttccctgtgtatcaatcttct 4560
Db 63906 ATTTTGATATTTGATTAATATATTTTCAATATCTATTTGTTTCTTTGTTTCAATATTT 63847
Qy 4561 atatatattgaacacatcttctcagaagagagagagagagagagagagagagagagagag 4620
Db 63846 ATATATTGAAGAAATCTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 63787
Qy 4621 catgcacacacagagagagagagagagagagagagagagagagagagagagagagagagag 4680
Db 63786 CATGCACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 63727
Qy 4681 tgcagaagctgaattagaagagagagagagagagagagagagagagagagagagagagag 4740
Db 63726 TGCAGAGACTGAATTAAGAAATTTCTCCCAAGATACAGAGTTGTTTAAAGCTAGGGGTG 63667
Qy 4741 aggggggaaatctgcagctctcagagagagagagagagagagagagagagagagagagag 4800
Db 63666 AGGGGGGAAATCTGCCCTTCTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 63607
Qy 4801 cttgtgtctgagctgtgtatcttctcgtctcgtctcgtctcgtctcgtctcgtctcgtct 4860
Db 63606 CTTGTGTCTGAGCTGTGCTGATTTTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 63547
Qy 4861 ggaatcccaagctctagacagagagagagagagagagagagagagagagagagagagagag 4920
Db 63546 GGATCTCCAGTTCCTACAGATAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 63487
Qy 4921 tgaatggaatatataaactagaatatatcttctgtctgaatcagaagagagagagagagag 4980
Db 63486 TGAATGGAATATTAATTAAGATATATCTCTGTTGAATTCAGCAGCAGCAGTACTCTGG 63427
Qy 4981 tgaatgagctgtgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5040
Db 63426 TGTAAAGTGTGTATC-----GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 63373
Qy 5041 gataaggaactatctatgtggagagagagagagagagagagagagagagagagagagagagag 5100
Db 63372 GATATAGGAATATTAATTTGGGTATGTGCTCATTAATTTGGAGATGTTCTTTTAAAGAA 63313
|||||
```

Qy 5101 actcaaacagactctcgtgaaggttatttcttaagaatctgtcgtcgaagcgtgaagcga 5160  
|||||  
Db 63312 ACTCCAAACAGACTCTGGAAGGTATTCTTAAGAAATTTCTGTGGCAGCGTGAAGGCA 63253  
Qy 5161 ccccccctgtcacaagccccccacagcctcagctggtccacctgtcttccccatgaag 5220  
|||||  
Db 63252 cccccctgtcacaagccccccacagcctcagctggtccacctgtcttccccatgaag 63193  
Qy 5221 gctggtctccccgttatataaactctcgtgaagcctggtgagccggaagcgaagcgaac 5280  
|||||  
Db 63192 GCTGCTCCCGAGTAAATGAACCTCTGTGAGCTGGGAGTGAAGCCAGCAAGCCAC 63133  
Qy 5281 catcagagcactcctcgaagcagc 5304  
|||||  
Db 63132 CATCAGGACCTCTCAGCAGC 63109

RESULT 2  
AF007562 5300 bp DNA PRI 18-MAR-1998  
LOCUS Homo sapiens trabecular meshwork inducible glucocorticoid response  
DEFINITION protein (TIGR) gene, promoter region and partial mRNA sequence.  
ACCESSION AF007562  
VERSION AF007562.1 GI:2970123  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 5300)  
AUTHORS Nguyen,T.D., Chen,P., Huang,W.D., Chen,H., Johnson,D. and  
Polansky,J.R.  
TITLE Gene structure and properties of TIGR, an olfactomedin-related  
glycoprotein cloned from glucocorticoid-induced trabecular meshwork  
cells  
JOURNAL J. Biol. Chem. 273 (11), 6341-6350 (1998)  
MEDLINE 98165818  
REFERENCE 2 (bases 1 to 5300)  
AUTHORS Nguyen,T.D., Chen,P., Chen,H. and Polansky,J.R.  
TITLE Direct Submission  
JOURNAL Submitted (10-JUN-1997) Ophthalmology, University of California San  
Francisco, 10 Kirkham Street, San Francisco, CA 94143-0730, USA  
LOCATION/Qualifiers

FEATURES  
source  
1..5300  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
gene  
1..>5300  
/gene="TIGR"  
promoter  
1..5271  
/gene="TIGR"  
5272..>5300  
mrna  
/gene="TIGR"  
/note="additional 3' sequence deposited as TIGR mRNA with  
Genbank Accession Number U85257"  
/product="trabecular meshwork inducible glucocorticoid  
response protein"

BASE COUNT 1483 a 1152 c 1236 g 1429 t  
ORIGIN

Query Match 99.0%; Score 5253.4; DB 11; Length 5300;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 5293; Conservative 0; Mismatches 6; Indels 6; Gaps 3;

Qy 1 actctgtcagttactcctcagggcctatatagaatgaataagataaccaaagtgaag 60  
|||||  
Db 1 ATCTTTGTCAGTTACTCTCAGGCGATATGAATGAATGAATGAATGAATGAATGAATGA 60  
Qy 61 tccataaacgtatagcctcattcgtatgtatgtcttgcgagatataagaatca 120  
|||||  
Db 61 TCCATTAACGTATAGCTCTCGATGTATGTCCTTTGGCAGGATGATTAAGATCA 120

Qy 121 ggaagaagagatataccagcttagcccaagtgctccaggtctgtctcttatttagtga 180  
|||||  
Db 121 GGAAGAAGAGATATCCAGCTTAGCCCAAGTGTGAGGCTGTGCTCTTATTATTAGTA 180  
Qy 181 cagatgtgtctctcgtcagaagcctattctcagaagaacataccataatgttaaac 240  
|||||  
Db 181 CAGATGTGTCTCTCAGAGAAAGCTATTCTTCAGAAATCATCATCATATAGTAATTC 240  
Qy 241 catcaaacagagagctcaagaagaatgaaga tgggcaacttgcacgaagaaatgccaag 300  
|||||  
Db 241 CATCAACAGAGAGCTCAAGAACGAAATGAGATGGGCACTTCCCAAGAAATGCGCAG 300  
Qy 301 gagaagcaataatgaataaataaacttctccctgtgtttaaattccaggaataatg 360  
|||||  
Db 301 GAGACCAATATATGATGAATAAATTAACCTTTGCTTTTATTTCAGGAAAAAATG 360  
Qy 361 atgaagaccataataatgaataaagaacagcccaagaaataatgtttcccaattgg 420  
|||||  
Db 361 ATGAGAGCCAAATCAATGAATGAAGAAACGCTCAGAAAAAAGATGTTCCAAATTGG 420  
Qy 421 taattaaatattgtctcttgggaagagaccccaatgtgagcttgatggaataatggaa 480  
|||||  
Db 421 TAATTAAATTTGTTCTCTTGGAAAGACACCTCCATGTGAGCTTATGGAATAATGGGA 480  
Qy 481 aaacgtcaaaagcagatcgtatcagatcccaagtggaattatatttaaaacacagat 540  
|||||  
Db 481 AAACGTCAAAAGCATGATCTGATCAGATCCCAAGTGATTTATTTAAAAACAGAT 540  
Qy 541 ggcatacctctggggagagcaagttcaagaaagtcattgttagcaaaagacatacaataac 600  
|||||  
Db 541 GCATCACCCTGGGAGGCAAGATTGAGAAAGTCAATGTAGCAAAAGACATACATATAC 600  
Qy 601 agcaaatcaaaaattccgcaaatgcaagaggaagaatgggagctgggaagccttcataac 660  
|||||  
Db 601 AGCAAAATCAAAATTCGCCAAATGCAGAGAAATGGGAGATGGGAATGGAATGTTATAC 660  
Qy 661 agtgaatagcagctgtgacatgttcgcaaacacccctccgtctataccagggaaacaaa 720  
|||||  
Db 661 AGTGAATTAGCGAGTTGACCATGTTCGCAACACCTCCGCTATACCGAGGAACAAAA 720  
Qy 721 attgactgggtcaaacctcggaacttcaaggggaataataaacctggaagcaaaacaaa 780  
|||||  
Db 721 ATTGACTGGGCTTAACCTCGGACTTTCAAGGAAATGAAATGAAATGAGAGCAAAACAAA 780  
Qy 781 gacatggtttaaaaggcaaacagacatgttgagccttcaaaagcagctgccccctcaaga 840  
|||||  
Db 781 GACATGTTAAAAAGCAACCAACATGTGAGCTTCAAAAGCAGACAGGCCCTCAGAGA 840  
Qy 841 gggaccctgagggcatttgcctttagaagagccaggttctcctaaggaaacttaagaactc 900  
|||||  
Db 841 GGGACCCTGAGGCATTTGCTTTAGGAAGGCCAGTTTCTTAAGCAATCTTAAGAAACTC 900  
Qy 901 ttgaagatcagaaatttaaacatttaagratataaacaataatgcatgataatgaag 960  
|||||  
Db 901 TTGAAGATCATGAATTTTAACATTTAAGATTAATAACAAATATGCGATGATATACG 960  
Qy 961 tttagacatgggtcccaatttataaagctcaagcattacaagagataacgtygtccagctcc 1020  
|||||  
Db 961 TTTAGACATGGTCCCATTTTATTAAGTCAAGCATACAGATTAACGATGATGATGATG 1020  
Qy 1021 ggaatagtcagaatcatatagaatcactgtgtcccatccttaacttctcagaatgac 1080  
|||||  
Db 1021 GGATAGTCAAGAAATCATTAAGAAATCACTGTGTGCTCCCATCTTAATCTTTTTCAGAAATGATC 1080  
Qy 1081 tgtcataagccctcaacacagagcccgatgtgtctgaactataacaaactctacaaccaa 1140  
|||||  
Db 1081 TGTCTATAGCCCTCACACACAGGCCGATGTGTACCTTACCAACACATCTAACAACCAA 1140  
Qy 1141 gtcctcaaacctgttgaagtgatcatctcagtaagttccattacaatgccaactccccc 1200  
|||||  
Db 1141 GTCCCTCAACCACTTTTAAACGTGATCTCAGTAGGTCCCATTAATAATGCCACCTCCCC 1200  
Qy 1201 tgtcagcccatcccgctccacaggaagttccccaacttgaagcttctgcatcagatgt 1260

[illegible]

Db	2281	CCCCAAGCCCGAGTCTTCCAAACCTCTCTTCATCACTACACAGCTGCAGCTGGCCT	2340
Qy	2341	gctctgcctcccgltgaatcgtctccgtgtgcatctgagcttgagacctctcttgctccagagct	2400
Db	2341	GCCTCGCTTCCCGGTAATCGTCTCGTGCATCTGAGCTTGAGACTCTTGGCTCCAGACT	2400
Qy	2401	CCGAAGAAGGAATtggagagaggaagaaactagctctaagcgagaatcttgaaggaggcagtgcttc	2460
Db	2401	CCAAAGGAATtGGAAGAGGGAACATAGCTTAACGAGAAATtGGAAGGAGACAGTGTTC	2460
Qy	2461	ctcagagggaaaggggagccctccacgctcccaagagaatctccagagagctggagagctgcaagggag	2520
Db	2461	CTCAGAGGGAAGAGGGGCTCTCAGCTCCAGAGAAATTCCAGAGAGTGGGAGCTGCAGGAG	2520
Qy	2521	tgggagacgcttgaggcttgagcgggtgtctgaagagcagaaagtgaagaaaggccaaggtcgaa	2580
Db	2521	TGGGAGAGCTGGGGGCTAGCGGGGCTGTGAAGCAGAGAGTGAAGAAAGGCAAGGCTGA	2580
Qy	2581	gctgcacagatgtctcagtgctgtctcaaggggctgggaagtttcggtctctccgttgggc	2640
Db	2581	GCTGCCAGAGTTCACATGTTGTTCAACGGGGGCTGGAGTGTTCCTTCCTCTGTGAGC	2640
Qy	2641	ctttatctctctctcgcgtctggagagaagaagctatcttcaatgaagagatgcagtttc	2700
Db	2641	CTTTTATCTTTCTCTCTGCTTGGAGAGAGAAGATCTATTCTATAGAGATGCAGTTTC	2700
Qy	2701	ataaagtcagctgttlaaatactcagggtgtgtcagctgggttttccctcagcgaagcccttat	2760
Db	2701	ATAAAGTCAGCTGTTAAATTTCCAGGGTGTGCATGGGTTCCTTCACGAAGCCTTAT	2760
Qy	2761	ttaatgggaatatatbgyaagcagagctcaattctcctagcgcttaattccaaggaagaaagtac	2820
Db	2761	TTAATGGGAATATGGAAGAGGAGACTCATTTCCTAGGCCCTTATATCAGGAAGATGAC	2820
Qy	2821	tggagctctctctctcaatgtctctctcgtgcaactcaactcaagccctgtgtgtggaacttgctta	2880
Db	2821	TGGAGCTCTTTCTTCTCATGTCTCTTGCGCACTACTACACCCTGTGTGTGACTTGGCTTA	2880
Qy	2881	tgcagaagcgtctcgaataaccccttgaatcagagaagatcgtgttctctctctgtgtctgccaat	2940
Db	2881	TGCAAGACGGTCCGAAMCCTTGGAATCAGAGACTGGTTCCTTCTTGAGTCTCCACTT	2940
Qy	2941	ggtctggtctgtgcgacccgttgggcaagtgctctctctctcccttgggccaatgctctctctgt	3000
Db	2941	GGTGTGCTGTGCGACCTGTGGCAAGTGTCTCTCTTCCCTGTGGCCATGTCTCTCTCT	3000
Qy	3001	ataaagacccctctgagctctcgtctgtctctgtgaacactctccctgtgtatctctctgtaaggg	3060
Db	3001	ATAAAGACCCCTTGACGCTCTCGTGTCTGTGAACACTTCCCTGTGTATCTCTGTAGGG	3060
Qy	3061	ggaatcttgaagagggaagagagcagagcttgaagcagcttgaaccacagggaggtgaaggg	3120
Db	3061	GGATGTTGAGAGGGGAGGAGGAGCGACGTGAGACACTGAGCCACAGGGGAGGTGAGGG	3120
Qy	3121	ggacaagaaaggcaggcagaaagctgggtgtctcactcaagttccactgtatcaagtcaagctc	3180
Db	3121	GGACAGGAAGCGACGAGAGCTGGGTGTCTCATCTCATCTCATCTCATCTCATCTC	3180
Qy	3181	caggaacggaagccaatactctcaggaagaagtcgaatgaaccacaagccacaatttct	3240
Db	3181	CAGGACCGAGGCGCCACATCTCTTCAGGAAGCTCAATGAACCAACGACGACATTTCT	3240
Qy	3241	tccctaagcatagacaattgcatcttgcacaaataacaaagaagatcagaagacttaactgt	3300
Db	3241	TCCCTAAGCATAGACATTTGCAATTTGCAATAACCAAAAGAAATGCAGAGACTACTGCT	3300
Qy	3301	ggtagcttctgtccttgcatctcaaaaacttggccagagcgaagtgaagaaatgcagagattg	3360
Db	3301	GGTAGCTTTCTTCGCGATTCAAAAACTGGGCGACAGCAAGTGAAGAAATGCACAGATTG	3360
Qy	3361	ttaaatcttcaaccttgacagacacccacacagagctcagcagtgactgtacagacagag	3420
Db	3361	TTAAACTTTTACCTTACCCACGACACCCACGCGCTCAGACAGTGACTGCTGTGACACAGG	3420



Oy	3421	agtgacccgcaagcgagggaggaagaaagagagagtagtgytgaagcaagaaag	3480
Db	3421	agtgacctgcagccagcgagggaggaagaaagagagagtagtgytgaagcaagaaag	3480
Oy	3461	acaggttcattcaagggcagcgagggaatttgaccacagggattatagccacgctgcgg	3540
Db	3461	acaggttcattcaagggcagcgagggaatttgaccacagggattatagccacgctgcgg	3540
Oy	3541	gttcgaagagcgagggcctatactctgggggaaaaaaatcagttcaggggaagtcggaga	3600
Db	3541	gttcgaagagcgagggcctatactctgggggaaaaaaatcagttcaggggaagtcggaga	3600
Oy	3601	ccgtgattcttaatacctabattttcccttcaagctgagtaattctgagcaagccaag	3660
Db	3601	cctgatttcttaatacctabattttcccttcaagctgagtaattctgagcaagccaag	3660
Oy	3661	gttagtaacttgagcgctgtgaaggttacctagtttcccttataggaaactttttctcgt	3720
Db	3661	gttagtaacttgagcgctgtgaaggttacctagtttcccttataggaaactttttctcgt	3720
Oy	3721	ggaggttcagcagcacaagggcacaatccgctttcttcaacaggaagaaacattcctaag	3780
Db	3721	ggaggttcagcagcacaagggcacaatccgctttcttcaacaggaagaaacattcctaag	3780
Oy	3781	taagccaacaagattccaagcctagctgtcgtgcactatagtatggttttttgaanaat	3840
Db	3781	taagccaacaagattccaagcctagctgtcgtgcactatagtatggttttttgaanaat	3840
Oy	3841	catctcagcagctgttcttactctgatactcagaanaatgagactgaactctgttcagctg	3900
Db	3841	catctcagcagctgttcttactctgatactcagaanaatgagactgaactctgttcagctg	3900
Oy	3901	taaacaaacaccccatctgttaaatgctcacaagctcaagcttaactgcagaaccaatcaat	3960
Db	3901	taaacaaacaccccatctgttaaatgctcacaagctcaagcttaactgcagaaccaatcaat	3960
Oy	3961	aagaataagatcctctagaagcaaatcgtgtttccac-ctcggaggtgagcttcccgagg	4019
Db	3961	aagaataagatcctctagaagcaaatcgtgtttccac-ctcggaggtgagcttcccgagg	4019
Oy	4020	cagcttggaatacttcttactccaagatctggaacccgttgttgtataaacaataag	4079
Db	4020	cagcttggaatacttcttactccaagatctggaacccgttgttgtataaacaataag	4079
Oy	4080	ttgcctcaaaagccaatcattatcttcaagtgtgcctaaagttactctgcagcttttgata	4139
Db	4080	ttgcctcaaaagccaatcattatcttcaagtgtgcctaaagttactctgcagcttttgata	4139
Oy	4140	tttatgtgctatctgcatttgccttttgccttttctcctcttggttcttaaatgataagca	4199
Db	4140	tttatgtgctatctgcatttgccttttgccttttctcctcttggttcttaaatgataagca	4199
Oy	4200	ggagattactaacccacagctccagaagaacctgtgaatttgaaagaagaaatactcgtt	4259
Db	4200	ggagattactaacccacagctccagaagaacctgtgaatttgaaagaagaaatactcgtt	4259
Oy	4260	ttatcttaccacaccttcaactaaattiaaactttatctccattgcgaatagaagccataa	4319
Db	4260	ttatcttaccacaccttcaactaaattiaaactttatctccattgcgaatagaagccataa	4319
Oy	4320	actcaaaagtgttaataagagtaacctcgtgatatttgcaataccaatgaaatacaagaat	4379
Db	4320	actcaaaagtgttaataagagtaacctcgtgatatttgcaataccaatgaaatacaagaat	4379
Oy	4380	tttatatacatactaacgttctgcaggttacgtgtgataagtgaaatatctactcaaaact	4439
Db	4380	tttatatacatactaacgttctgcaggttacgtgtgataagtgaaatatctactcaaaact	4439
Oy	4440	actttgaaattagaacctcctgcgtggaactcgttttcaacatactaaataaactgtttaa	4499
Db	4440	actttgaaattagaacctcctgcgtggaactcgttttcaacatactaaataaactgtttaa	4499

[illegible]



REFERENCE  
AUTHORS

2 (bases 1 to 170425)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Beda, F., Boguski, L.,  
 Boucknight, B., Brown, A., Burkett, C., Campilano, A., Castle, A.,  
 Choe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
 Deaellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M.,  
 Fears, J., Ferreira, P., Fitzhugh, M., Forrest, C., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Grant, G., Hags, B., Heaford, A., Horton, L.,  
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kenn, L., Karats, A.,  
 Klein, J., Landers, T., Largocque, K., Lehoczy, J., Levine, R.,  
 Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M.,  
 McEwan, P., McGuirk, A., McKernan, K., McPheters, R., Meldrum, J.,  
 Menus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Naylor, J.,  
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M.,  
 Peterson, K., Pierre, N., Pisenl, C., Pollara, V., Raymond, C.,  
 Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,  
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Subramanian, A., Talamas, J., Teafaye, S., Theodore, J., Tirrell, A.,  
 Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,  
 Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and  
 Zody, M.

Direct Submission  
 Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On May 25, 2000 this sequence version replaced gi:7249345.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996:1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
 Center project name: L7153  
 Center clone name: 138\_F\_3

----- Summary Statistics  
 Sequencing vector: M13, M7815, 100% of reads  
 Chemistry: Dye-terminator Big Dye, 100% of reads  
 Assembly program: Phrap, version 0.960731  
 Consensus quality: 151483 bases at least Q40  
 Consensus quality: 160439 bases at least Q30  
 Consensus quality: 164708 bases at least Q20  
 Insert size: 16700; agarose-ef  
 Insert size: 167725; sum-of-contents  
 Quality coverage: 4.1 in Q20 bases; agarose-ef  
 Quality coverage: 4.1 in Q20 bases; sum-of-contents

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 28 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1  
 1301 1400: contig of 1300 bp in length  
 1401 2412: contig of 1012 bp in length  
 2413 2512: gap of 100 bp  
 2513 4147: contig of 1635 bp in length  
 4148 4247: gap of 100 bp  
 4248 5455: contig of 1208 bp in length  
 5456 5555: gap of 100 bp  
 5556 6997: contig of 1442 bp in length  
 6998 7097: gap of 100 bp  
 7098 8534: contig of 1437 bp in length  
 8535 8634: gap of 100 bp  
 8635 10164: contig of 1530 bp in length  
 10165 10264: gap of 100 bp  
 10265 10487: contig of 223 bp in length  
 10488 10587: gap of 100 bp

\* 10588 12420: contig of 1833 bp in length  
 \* 12421 12520: gap of 100 bp  
 \* 12521 14843: contig of 2323 bp in length  
 \* 14844 14943: gap of 100 bp  
 \* 14944 17265: contig of 2322 bp in length  
 \* 17266 17365: gap of 100 bp  
 \* 17366 20071: contig of 2706 bp in length  
 \* 20072 20171: gap of 100 bp  
 \* 20172 23531: contig of 3360 bp in length  
 \* 23532 23631: gap of 100 bp  
 \* 23632 26137: contig of 2506 bp in length  
 \* 26138 26237: gap of 100 bp  
 \* 26238 29858: contig of 3621 bp in length  
 \* 29859 29958: gap of 100 bp  
 \* 29959 36225: contig of 6267 bp in length  
 \* 36226 36325: gap of 100 bp  
 \* 36326 44015: contig of 7690 bp in length  
 \* 44016 44115: gap of 100 bp  
 \* 44116 49493: contig of 5378 bp in length  
 \* 49494 49593: gap of 100 bp  
 \* 49594 56796: contig of 7203 bp in length  
 \* 56797 56896: gap of 100 bp  
 \* 56897 62860: contig of 5964 bp in length  
 \* 62861 62960: gap of 100 bp  
 \* 62961 72552: contig of 9592 bp in length  
 \* 72553 72652: gap of 100 bp  
 \* 72653 83310: contig of 10658 bp in length  
 \* 83311 83410: gap of 100 bp  
 \* 83411 92992: contig of 9582 bp in length  
 \* 92993 93092: gap of 100 bp  
 \* 93093 103886: contig of 10794 bp in length  
 \* 103887 103986: gap of 100 bp  
 \* 103987 119418: contig of 15432 bp in length  
 \* 119419 119518: gap of 100 bp  
 \* 119519 136955: contig of 17177 bp in length  
 \* 136956 136795: gap of 100 bp  
 \* 136796 153977: contig of 17182 bp in length  
 \* 153978 154077: gap of 100 bp  
 \* 154078 170425: contig of 16548 bp in length.

FEATURES  
 source  
 1. 170425  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /map="1"  
 /clone="RP11-138F3"  
 /clone\_11b="RP11 Human Male BAC"  
 1. 1300  
 /note="assembly\_fragment"  
 misc\_feature  
 1401. 2412  
 /note="assembly\_fragment"  
 misc\_feature  
 2513. 4147  
 /note="assembly\_fragment"  
 misc\_feature  
 4248. 5455  
 /note="assembly\_fragment"  
 misc\_feature  
 5556. 6997  
 /note="assembly\_fragment"  
 misc\_feature  
 7098. 8534  
 /note="assembly\_fragment"  
 misc\_feature  
 8635. 10164  
 /note="assembly\_fragment"  
 misc\_feature  
 10265. 10487  
 /note="assembly\_fragment"  
 misc\_feature  
 clone\_end:SP6  
 vector\_side:right  
 10588. 12420  
 /note="assembly\_fragment"  
 misc\_feature  
 12521. 14843  
 /note="assembly\_fragment"  
 misc\_feature  
 14944. 17265  
 /note="assembly\_fragment"  
 misc\_feature  
 17366. 20071  
 /note="assembly\_fragment"



Oy	1556	agtgcttaataagaataaactctgaatggttacttaaaccaacaggaataacagcaaaagc	1615
Db	65245	ACTGTTTATAATAGGATATACTTGAATGGTTTACTAAACCAACAGGGAAACAGCAAAAGC	65304
Oy	1616	tgtgaataattcaaggagaattcttggggatctggggaaatggctgcataagcttgccctgactgccc	1675
Db	65305	TGTGATTAATTTTACGGGATTCTTGGGATGGGGAAATGGTGGCATGAGCTGGCTGGCTTGTGC	65366
Oy	1676	cagaccacttgctcctcaactctctccctccatccatcttcaaggtcaagttaaccatt	1735
Db	65365	CAGACCACACTGGTCTCATCATCTTCTTCCCTCATCTCTCAATTTTCAGGCTAAGTTACCATT	65422
Oy	1736	ttaatcaaccaatgctcttctgtgttaagcctccacatcgcttacttgaataaagatataata	1795
Db	65425	TTATTCACCATCTCTTGTGGTAAGGCTCCACATCCCTACTGTAATAAGAGTTAATAATA	65484
Oy	1796	actagttccactctggggccactctgtgtgtgtatagggagaggagcaataccccaagag	1855
Db	65485	ACTAGTTCCATTTTGGGGCCATCTGTGTGTGTATAGGGAGAGGAGCATACCCCAAGAG	65544
Oy	1856	ctcctctgaagcccccgcgcagaaagtttctctctccagcttggggagccctgcaagcaccgcg	1915
Db	65545	CTCCTTGAAGCCCCCGCGCAGAGGTTTCCTCCACGCTGGGGAGCCCTGCACACACCGG	65604
Oy	1916	ggtcctgggtgtgtccctgagaacaactgcgaacccgtgcacatggtgtgttcttcaactctc	1975
Db	65605	GGTCTCTGGGTGTCTCTGAGCAACCTGCCACCCCTGCCACAGTGTGTTGTATCACTCT	65666
Oy	1976	ctaggaagcctgtgtctcttactctctgtgttgaactgtgttcaatcaacagatcataatga	2035
Db	65665	CTAGGGAACCTGTGCTTCTTATTTTCTGTGACTCGTTCAATTATTCAGGCATTTCATTGA	65722
Oy	2036	caatttatgaatlaactatatactgcccagacaaccagagacaanaatgtgtgaagcaagcagtc	2095
Db	65725	CAATTATATGAGACTTATATATCTGCAGACACCAAGACAAATATGGAGCAAAAGCAGCTC	65784
Oy	2096	actgcctcaactctctgtggaggtgtgaacagttctcataatgaagaagctgcagaagaataat	2155
Db	65785	ACTGCCCTACCTTCGTGGGAGGTGACAGTTCTCATATGAAGACCTGCAGAAAGAAATTAAT	65844
Oy	2156	agccagagccaacttaaaccccaagtcgtcaaaagaaaggaataaacaacatcttgaagaatg	2215
Db	65845	AGCCAGCCACTTAAACCAAGTGTCTAAAAGAAAGAAATTAACACCATCTTGAAGAATTG	65904
Oy	2216	tgcgcagcaatcccttcaacaaggccacatccctctagcgccccctgtgcctccatcgttgccc	2275
Db	65905	TGCGCAGATCCCTTAAACAAGGCCACTCCCTAGCGCCCCCTGCTGCTCATCTGATGCGCC	65966
Oy	2276	ggaagcccccaagcccgagttcttcaagcctctctctccatcaagcaagcgtctgaagct	2335
Db	65965	GGAGGCCCCCAAGCCCGAGTCTTCCAAAGCCTCTCTCTCCATCAAGTACACGCTGAGACT	66022
Oy	2336	ggccctgcctccgacttcccgcgtgaatccgtccctcgtgtgcattctgaagtcgtggagaccccttgctcc	2395
Db	66025	GGCCTGCGCTCTCCCGTGAATCGTCTGTGTGATCTGTGACTGTGAGACTCTTGGCTCC	66084
Oy	2396	aggctccagaaagaaatgtgaaggggaaactagctcaacaggaagatctggagggagacgt	2455
Db	66085	AGGCTCCAAAGAAATGTGAGAGGGAACACTATCTTAACGGAATACTGAGAGGGAGACT	66144
Oy	2456	gttctccagaggaagaaaggccctccagctccagaggaattccaggaagtgtgggaactgca	2515
Db	66145	GTTTCTCTCAGAGGAAGAGGGCCTCCACAGCTCCAGGAATATCCAGAGAGGTGGGACTGCA	66202
Oy	2516	gggaagtgtgggaagcctgtgggctctgaagcggtgtgtctaaaggcagaggaagtgtaaaaggcaag	2575
Db	66205	GGGAGTGGGAGCGCTGGGCTGAGCGGTGTCTAAAAGCCAGAGGTGAAGAAAGGCCAAGG	66266
Oy	2576	ctgaagactgcacacagatcttccacagtgtgtgttcaacggggcttgggaattcttcgcttgcctg	2635
Db	66265	CTGAAAGCTCCCAAGATGTTCAAGTGTGTTCACGGGCTGGGAATTTTTCCTTGTCTTCTCTG	66322
Oy	2636	tgaagcctttatcttcttctcgtcttggagagaaagatctatctcaatgaaggaatgca	2695

Dd	66325	TGACCCTTTTATTCTTTTCTCTGCTTGGAGGAGAAAGTATTATTATGTAAGGATGCA	66384
Qy	2696	gttccataaagtcacgtctgttaaatccaaaggtctgtcaatgtttcttcctcaacgaagcc	2755
Dd	66385	GTTTCATAAAGTCAGCGTTTAAATTTCCAGGGTGTCATGGGTTTCTTCCACGAAGCC	66444
Qy	2756	ttaattcaatggaatatagaagcagagctcaattctcaagccgttaataccaagaa	2815
Dd	66445	TTTATTTAAGGAAATATAGAAAGCAGCTCATTTCTTAGCCCTTATTCACGAGAA	66504
Qy	2816	gtgactggaattcttcttctcaatgtctctgtggaactaccccaacccctgtgtgagcttg	2875
Dd	66505	GTGACTGGAAGCTTTTCTTTCACTGCTTCTGGGCAACTACACCCCTGTGTGTGACTTG	66566
Qy	2876	gcttatccaagacgcttcgaaaacctctggaatacagaagacatcggtcttcttctgtcttg	2935
Dd	66565	GCTTATTCGAAGACGGTCGAAAACTTGGAAATAGAGACATGGGTTTCTTTCTGTGTTCTG	66624
Qy	2936	ccattgtctgtctgtgtcgacgctgtggacaagtctctctctccctgtggccaatgctctc	2995
Dd	66625	CCATTCTTTGGCTGTGTGACCGGTGGGCAAGTGTCTCTCTTCCCTGGGCAATATCTTCT	66688
Qy	2996	ctgtactaaagaccctctgcaagcctcgtgtgtctctgtgaaacattccctgtatctctgtg	3055
Dd	66685	CTGCTATTAAGACCTTGTGACGCTCTGTGTGTGTGTGAACACTTCCCTGTGTCTCTGTG	66744
Qy	3056	agggggagatctgaagagggaaagagacgagactgtgaagcaactgtgacacacagggaggtg	3115
Dd	66745	AGGGGGGATGTGAGAGGGGAAGAGGACGAGACTGTGACACTGTGACGACGACAGCGGGAAGTG	66800
Qy	3116	gaaggaggacaggaagcgaagcagaagctgtgtgtctcaactcaagctccactgatacagta	3175
Dd	66805	GAGGGGACACGGAAGCGAGGCAAAACCTGTGGTGCTCATCATGCTCATCATGATCATCTCA	66866
Qy	3176	gacccacagagccgaagagcccaatgtcttaagaagaagctcaatgaaacccaacagccaatc	3235
Dd	66865	GACTCCAGAGCCGAGAGCCCAATGCTTGAAGAAAGCTCAATGAAACCCACACGACCAATT	66922
Qy	3236	tctctccctcaagacatagaacaatgtgcatctgtgcaataaaccaaaaaagatgcagaagataa	3295
Dd	66925	TTCTCTCCCTAAGATATGACAAATGGCATTTGGCATAATCAAAAAAGATGAGAGACTAA	66988
Qy	3296	ctgtgtgtatgacttctgtcctgtgcaattcaaaaactgtggccagagcaagtgtgaataatgcaga	3355
Dd	66985	CTGTGTGTACTTTTGTGCTGTGCAATTTCAAAACTGTGGCCAGAGCAAGTGGAATAAGCCGCA	67044
Qy	3356	gattgttaaacttttaacctgtgacccgaccccccaagagcgtccagcagatgtgactgtgtgacg	3415
Dd	67045	GATTGTAAACTTTTCACTCTGTGACGACGCCACCCACAGCAAGCTCAGAGTACTGTGTGACG	67100
Qy	3416	caaggagtgaccctgcagcagcagggggagggagaaagaaagagagagatgtatlatgagcaaa	3475
Dd	67105	CACGAGTAGACCTGTGACGCGAGGGGAGGAGAAAGAAAGAGGGGATATGTATATGACAA	67166
Qy	3476	gaagacagatatcaattcaaggcagatggaattgtgacacagggatlatagtlcaccgtgat	3535
Dd	67165	GAAAGACAGATTCAATTCAAGGCGCAGTGGGAATTTGACCAAGGGATTATAGTCCACGTGAT	67224
Qy	3536	ccgtgtgtctcaaggagcgaagggcctatatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	3595
Dd	67225	CCTGGGTTTCAAGGAGCGAGGCGTATTTTGTGGGGGAAAAAATCAGTTCAAGGGAAATCG	67288
Qy	3596	ggaagaccgtgattctctaatactatattcttccttcaacagcgtgaatlatcttgagcaagtc	3655
Dd	67285	GGAAGACTGATTCTTAATTAATATTTCCTTTTCAAGGCTGAGTAATTCAGAACAGTC	67344
Qy	3656	acaaggtagtaactbgagcgtgtataagattactagttcttccattataggaactcttttc	3715
Dd	67345	ACAAGGTAGTAAGTGAAGGCGTGAAGATTAAGTTTCCCTTATTAAGAACTCTTTTTC	67400
Qy	3716	tctgtgagatgaagacaaagggcaatcccgattctctttaaaggaaagaaacattcct	3775

Dh	67405	TCGTGCGAGTTAGCAGACACAAAGGCCAATCCGTTTCCTTTTAACAGAGAAAAACATTCCT	67468
Oy	3776	aagagtaaaagccaaacagaatccaagctcgaagctctctgcatacatabatgattglttctga	3835
Dh	67465	AAGAGTAAGGCCAAACAGATCTCAAGCGCTGAGTCTCTGACTATATGATGGTTTTTGA	67524
Oy	3836	aaaaatcaattcagcgatgtttacatactgaatccaagaataatgagactagtaaccttgc	3895
Dh	67525	AAAATCATTTTCAGCGATGTTTACTATCTATCTATTACAGAAATGAGACATGACCTTTGGTC	67584
Oy	3896	agccttcaaaccaaacaccactcttgaagttccaaagtctgaagcttbaagctbaacgcgaacaa	3955
Dh	67585	ACGTGTAACCAACACACCCTTTGTAAATCTCTCAAGTTAGGCTTACTGCGAACCAAT	67644
Oy	3956	caaaataagaataaactctttaagccaacatgtgttctcacctctgaaggtgagctgc	4015
Dh	67645	CAAAATAGAAATAGAAATCTTTAAGCAAACTGATGTTTCTCCACTCTGAGAGTGCTGCC	67704
Oy	4016	agggcagctctgaaataatttacttccaaagtattgaacctgtgtgtgtaataacaact	4075
Dh	67705	AGGGCAGTTTGAAATATTACTTCCAAAGATATGACATGTTGTGTGATTAACAAAT	67764
Oy	4076	aaagtctccaaaggccaatcatatttcaagtcggcctaaagttaactctgcagatttgg	4135
Dh	67765	AAAGTTGCTCAAGGCAATCATATTTCAGTGGCTTAAGTTACTCTGACAGTTTGG	67824
Oy	4136	tataatcaatggcattggccattggcttcttcttctctcttcccttggattaatgta	4195
Dh	67825	TATATTATTATGCGTATGCGCATTTGCTTTTGTGTTTCTCTTGGGTTATTAAATGTA	67884
Oy	4196	agcagggatattaaacctcagttccagaagcctgtgaatttgaatgagaaaaatttc	4255
Dh	67885	ACCAAGGATATTAACTAGACGTCCAGAAAGCCTGTGAATTTGAAGAGAAAAATTAC	67944
Oy	4256	gtttctatttttcaacaccttcaactcaaaatttcaacttatacttcactgcgaatagacc	4315
Dh	67945	ATTTTGTGTTTACCACTCTTCTACTAAATTTAACATTTTATTCATTTCGATAGACC	68004
Oy	4316	ataaacctcaagtcggtlaaataagaatacctctgatttctgcatccaalagaatacacg	4375
Dh	68005	ATTAACCTCAAACTGGTAATTAACAGTACCTGATTTTGTCATTACCAATAGAATCACAG	68064
Oy	4376	aaatttataccatatttaacagttgttgcgaagtgaagtgtgaagtgaataattatactaa	4435
Dh	68065	ACATTTTAATACATATTAACAGTGTGTGCAATACGTTGAATGGAATTTTATACTCA	68124
Oy	4436	aacctacttgaattagaccctcctgcgtgaccttgtttttaaataataaataacgt	4495
Dh	68125	AACTACTTTGAATTAAGACCTCTGCTGATCTGTGTTTAACTATTAATTAACATGT	68184
Oy	4496	ttaaaatttgaataatttgaataacataattcatatactattgtttccttcttgaatcat	4555
Dh	68185	TTAAATTTTGATATTTTGATATCATATTTTCATTATCTATTGTTTCTCTTGTAACTAT	68244
Oy	4556	atttatataatttgaaaaaaacatcttcttgaagaaggttccccaagatttcaaccaatgagt	4615
Dh	68245	ATTTTATATATTGAAAAACATCTTTCTGAAAGAGTTCGCCAGATTTCACCAATGAGGT	68304
Oy	4616	cttggcagtcacacacacagagtaagaacttgaattgaagagctbaactgaacttgc	4675
Dh	68305	CTTGCACTGACACACACAGATTAAGACTGATTTTGAAGGCTTAACATTGAGCTG	68364
Oy	4676	tgaagtcgaagagctgaatttgaagaagcttcccaagaatacacagctgtttttaaagctag	4735
Dh	68365	TGAGATGCAAGACTGAATTAAGAAGTCTCCCAAGAATACACAGTGTGTTTAAAGCTAG	68424
Oy	4736	gggtgaagggggaaatctgcgccttcataaggaatgtctccctggagcctgtgaagtg	4795
Dh	68425	GGGTGAGGGGGGAAATCTCCCTCTTAAGGAATGCTCTCCCTGGAGCCTGTAAGGCTG	68484
Oy	4796	ctgtccctctgtgtctcggcgcgcgtgtatttttctctgttccctgcgtcaagctctaaggact	4855
Dh	68485	CTGTCTCTGTGTCTTGCGCTGCTGTATTTTCTGTCTCCGCTACACGCTTAAAGGACT	68544

[illegible]

RESULT	4
LSMTCOC1	
LOCUS	HSMTYOC1      2800 bp      DNA      PRI      21-APR-1998
DEFINITION	Homo sapiens myocillin (GLC1A) gene, promoter and exon 1.
ACCESSION	AF049791
VERSION	AF049791.1    GI:3065670
KEYWORDS	
SEGMENT	1 of 3
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia; Eutheria: Primates: Catarrhini: Homnidae: Homo. 1 (bases 1 to 2800)
AUTHORS	Fingert,J.H., Ying,L., Swiderski,R.E., Nystuen,A.M., Arbour,N.C., Alward,W.L.M., Sheffield,V.C. and Stone,E.M. Characterization and comparison of the human and mouse GLC1A glaucoma genes
TITLE	Genome Res. (1998). In press 2 (bases 1 to 2800) Fingert,J.H., Ying,L., Swiderski,R.E., Nystuen,A.M., Arbour,N.C., Alward,W.L.M., Sheffield,V.C. and Stone,E.M. Direct Submission
JOURNAL	Submitted (23-FEB-1998) Ophthalmology, University of Iowa, 200 Hawkins Drive, Iowa City, IA 52242, USA
FEATURES	Location/Qualifiers
source	1..2800
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="1"
	/map="1q21-q31"
	1..1905
promoter	/gene="GLC1A"
exon	1906..2509
	/gene="GLC1A"
	/number=1
BASE COUNT	781 a      588 c      673 g      758 t
ORIGIN	

Query Match	34.6%	Score 1936	DB 66	Length 2800	
Best Local Similarity	99.5%	Pred. No. 0			
Matches 1864	Conservative	0	Mismatches	5	Indels
				5	Gaps
Oy	3431	agcgcagggaggaagaaagaaagagagatagtcgtatgagcaagaagacatatcat	3490		
Db	1	AGCGCAGGGGAGGAGAGAG-AAAAGAGAGGGAGATAGTGTATGACCAAGAAAGACAGATTCA	59		
Oy	3491	tcaagggcagtagtggaaattgacacacagggaattatagtcacgctgatactctggtcttagag	3550		
Db	60	TCAAGGCGCAGTGGGAATTGACCACAGGGATATATAGTCCAGTATCCTGGTTCTAGAG	119		
Oy	3551	gcagggcataattgtgvgggggaaaaaaacagttcaagggaagtcgggagacctgattct	3610		
Db	120	GCAGGCGTATATTTGGGGGGAAAAAATACGTTCAAGGGAAAGTCCGGACCTGATTCT	179		
Oy	3611	aatactatactttctctcttaccagcagtgtaattctcgagcaagtcacaggtcgaactg	3670		
Db	180	AATACTATATTTTCTCTTTACAACTGACTAATTTCTGACCAAGTCCANAGTATACTG	239		
Oy	3671	aggcgtcaagaattacttagttctctcttacttaggaactctttctctgttggagttaga	3730		
Db	240	AGCGTGAAGATTAAGTACTGATTCTCTCTTATPAGAACTCTTTTCTCTGTGAGTAGCA	299		
Oy	3731	gcacaaagggcaatcccgcttctctttaaacaggaagaacaatccctaagaagttaaagcaa	3790		
Db	300	GCACAGAGGCAATCCCGTTTCTTTAACCAGGAAGAAACATTTCTPAGAGTAAAGCCAA	359		
Oy	3791	cagatccaagccctaaagtcctgtctgtactataatgattggtttctgaaaaatcatctcag	3850		
Db	360	CAGATTCAAGCTAGGCTGTCTGTGATATATGATGTGTTTGAAAAATCATTTTCAGG	419		
Oy	3851	atgttctactatctgattcagaaatagagactagtaacctgttgcagctgtataacaaca	3910		
Db	420	ATGTTACTATGTGATTCGAAATAGAGACTGTACCCTTGGTACGTGTAACAAACA	479		
Oy	3911	cccatctgtaaatgctctcaagttcaggcttcaactgcagaaaccaatcaataaagaaataga	3970		
Db	480	CCCATTTGTAAATGCTCTCAAGTTCAAGGCTTAAGTCCAGAACCAATCCAAATAAGAAATGAA	539		
Oy	3971	tctttagagcaaacgtgtcttccactctgtgaagtgagtcgcgcaggcagtttggaaa	4030		
Db	540	TCTTAGACCAACAGCTGTGTCTCCACTCTGGAGGTGAGCTGTGCCAGGCGACTGTGAAA	599		
Oy	4031	tatttactccaagtiattgacacgtgtgtgtatctaaacaataaagttgctcaag	4090		
Db	600	TATTACTTTCACAAATATTGACACGTGTGTGTGATATTAAACAACATATAATGCTCAAG	659		
Oy	4091	caatcatacttccaagtgvgcttaaagttactctgcagcttctgtatattatgtgcta	4150		
Db	660	CAATCATTAATTCAAGTGGCTTAAAGTTACTGTGACAGTTTGGTATATTATTTGGCTA	719		
Oy	4151	tgtcgcaattgcttctgttttctctctcttgggttttaatagttaaagcgggattcttaa	4210		
Db	720	TTGCAATTTGGCTTTTGTGTTTCTCTTTGGGTTTATATATGTAAGACGAGGATTAATTAA	779		
Oy	4211	ccctacagtcagaaagccgtgtgaatttgaatgagagaaaaatagcttttattcttacc	4270		
Db	780	CCTACAGTCCAAAGAACCTGTGAATTTGAATGAGGAAAAAATTACTTTTGTGTTTAA	839		
Oy	4271	acctctcaactaaatttaacattttatccatttgcgaatagagccaataactcaaaigtg	4330		
Db	840	ACCTTCTACTAATTAATTAACTTTATTCATTTGCAATAGAGCCATAAATCAAGTGG	899		
Oy	4331	taataaagttacctgtgattctgtcatctacaaatgaaatcaacagcatcttactctata	4390		
Db	900	TATATACATFACCTGTGATTGTGTCTATTCCAAATATAAATTCACAGCAATTTATATATA	959		
Oy	4391	ttaacagttgtcaggtacagctgttaagtgaaatatctacatcaaaacacttggaaatt	4450		
Db	960	TTACAGTTTGTTCAGATACGATACGTTGATAGTAAATATTATTATCTCAAAACTACTTTGAAAT	1019		
Oy	4451	agacctccgcctgtagtctgttttcaacatatbaataaaaatgltttaaattttgatat	4510		

Db	1020	AGACCTCTGCTGGATCTTTTAAACATTAATAAACAATGTTTAAATTTTGATAT	1079
Qy	4511	cttgataatacaatcctcatcatcattcgttcctctgtatactcaatctcatalatctga	4570
Db	1080	TTTATATATCATATTTCATATATCATTTGTTTCTTGTGTAATCATATATTATATTGGA	1139
Qy	4571	aaacacctcttcctgaagaagctccccaagatctcccaalgaagctctctggcaatcaca	4630
Db	1140	AAACATCTTCTCGAAGAGTTCCTCCAGATTCACCAATGAGAGTTCCTGGATGCACACA	1199
Qy	4631	cacagagtaagaactgaattagaagctaaatctgaatctggtccttgagaatgaagactg	4690
Db	1200	CACAGAGTAAGAATGTATTTAGAGGCTAACATTGAATGATGGCTGGATGACAAGACTG	1259
Qy	4691	aaatagaagaagctctcccaagaatacacaagctgttttaaagctaaaggctgaaggaggaa	4750
Db	1260	AAATTAAGAAGTTCTCCCAAGATACACAGTTGTTTAAAGCTAAGGCTGAGGGGAGAA	1319
Qy	4751	tcgcgcgccttcataagaatgcctccctggagccctgtlaaggcgtcctcgtgtctc	4810
Db	1320	TCTGCCCTTCTATAGGAATGCTCTCCCTGGAGCCGTGATGGGTCTCTTGTTGTTCT	1379
Qy	4811	ggcgcgcgttatattctctctgcctctgaagctcttaaggaaactgtttgatactcag	4870
Db	1380	GGCGCGCTGATATTTCTGTGCTCCGTACGCTTTAAAGGACTGTTGGATCTCCAG	1439
Qy	4871	cttcctagcatagtcgctgagcagctgcagcttcctcaatgaatttcagaagttaatgaa	4930
Db	1440	TTCTCAGCATATGTCCTCCGACAGTCGAGTTCTCAATAGTTGCAGAGTAATGAGAA	1499
Qy	4931	tataactagaabaatalactctgtctgaaatcagcacaccagtaagctcgtgtgtaagtg	4990
Db	1500	TATAAAGTAAGAATATATCTCTGTGAATAACAGCACACAGATGCTGGTGTAGTGTG	1559
Qy	4991	tgtatcggt	5050
Db	1560	TGTAC---GtGtGTGtGTGTGTGTGTGTGTGTGTAAACAGAGTGAAGATATAGAA	1615
Qy	5051	ctatatatggggatagtggtgcataaatctggatgtctcttttaaagaagaactccaaca	5110
Db	1616	CTATATATGGGGTATGGGTGCATTAATTTGGAGATGTTCTTTTAAAGAAACTCCAAACA	1675
Qy	5111	gactctcgtgaaggttattcttctaagaatctgcgtgcagcgtgaaggcaaacccccctgt	5170
Db	1676	GACTTCTGGAAGGTTATTTCTTAAGAATCTGTCTGACACGTGAAGGCAACCCCTGTG	1735
Qy	5171	cacagccccaccagcctcagctggtgcacactgtcttctcccaatgaagggtgtgtctcc	5230
Db	1736	CACAGCCCCACCCGCTCTCAGTGGCCACTCTGTCTTCCCAATGAAGGGCTGTGCTCC	1795
Qy	5231	cagatataataaacctctctggagctcgggcatgaagcagaagaagccaccccatccaggca	5290
Db	1796	CAGTATATATTAACCTTCTTGAGAGCTTGCGCATGAGCCAGCAAGGCCATCTCAGGCA	1855
Qy	5291	cctctcagcaagc 5304	
Db	1856	CCTCTCAGCAAGC 1869	
RESULT	5		
AB006686S1			
LOCUS	AB006686s1	1228 bp	DNA
DEFINITION	Homo sapiens gene for myocillin, exon 1, complete sequence.	PRI	14-Apr-2000
ACCESSION	AB006686		
VERSION	AB006686.1	GI:2828342	
KEYWORDS	myocillin.		
SEGMENT	1 of 3		
SOURCE	Homo sapiens pre-pro-B cell clone_KB180C12.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		

[illegible]

QY	5042	attagagacattatgtgggtgttggtgataaatctggatgtctcttttaaaaaagaa	5101
Db	175	ATATGAGACATTAATTATGGGTATGGGTGATTAATTTGGATGTTCTTTTAAAAAGAA	234
QY	5102	ctccaaagaagctcttgaaggtattcttctaaagaatctctgtgcagcgltgaaggaac	5161
Db	235	CTCCAAAGACACTTCGGGAAGGTATTCTTGTAGAACTTCGTGGGAGCGTGAAGCAAC	294
QY	5162	ccccctgtcaagagcccccacagcctcaagctgagcactctgtcttcccccataaagg	5221
Db	295	CCCCCTGTGCGACACCCACCAGCCCTCACAGTGGCCACCCTCTCTCTCCCATGAAGG	354
QY	5222	ctggtctcccccagatataataacctctctgaagctcgtggcatgagccagcaagccacc	5281
Db	355	CTGGCTCCCCAGTATTAATTAACCTCTCTGAGACTGGGCGATGAGCCAGCAAGGCCACC	414
QY	5282	atccagagcactctcagaacagc	5304
Db	415	ATCCAGGCACCTCTCAGCACAGC	437
RESULT	6		
HSGLC1A1			
LOCUS	1086 bp	DNA	PRI
DEFINITION	Homo sapiens GLC1A (trabecular meshwork induced glucocorticoid response) gene, exon 1, joined CDS.		30-JAN-1998
ACCESSION	297171		
VERSION	297171.1	GI:2425156	
KEYWORDS	GLC1A.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1086)		
AUTHORS	Stone,E.N., Fingerhut,J.H., Alward,W.L., Nguyen,T.D., Polansky,J.R., Sundén,S.L., Nishimura,D., Clark,A.F., Nystuen,A., Nichols,B.E., Ritch,R., Kalenak,J.W., Craven,E.R. and Sheffield,V.C.		
TITLE	Identification of a gene that causes primary open angle glaucoma		
JOURNAL	Science 275 (5300), 668-670 (1997)		
MEDLINE	97158493		
REFERENCE	2 (bases 1 to 1086)		
AUTHORS	Adam,M.F., Belmouden,A., Binisti,P., Brezin,A.P., Valtot,F., Bechevillie,A., Descotte,J.C., Copin,B., Gomez,L., Chaventre,A., Bach,J.F. and Garchon,H.J.		
TITLE	Recurrent mutations in a single exon encoding the evolutionarily conserved olfactomedin-homology domain of TIGR in familial open angle glaucoma		
JOURNAL	Hum. Mol. Genet. 6 (12), 2091-2097 (1997)		
MEDLINE	97472461		
REFERENCE	3 (bases 1 to 1086)		
AUTHORS	Garchon,H.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-JUN-1997) Garchon H.-J., Hopital Necker, INSERM U25, 161 rue de Sevres 75743 Paris cedex 15 FRANCE		
FEATURES	Location/Qualifiers		
source	1..1086		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="q23-q25"		
	/tissue_type="white blood cells"		
TATA_signal	355..363		
exon	385..1065		
	/gene="GLC1A"		
	/label=ex1		
	/number=1		
MRNA	join(385..1065,297177.1:21..146,297174.1:21..1475)		
	/gene="RNGLC1A"		
	/label="RNGLC1A"		
gene	join(385..1065,297177.1:21..146,297174.1:21..1475)		
	/gene="GLC1A"		
misc_feature	462..1065		

```

/gene="GLC1A"
/note="first coding exon"
/label=start
join(462..1065,297177.1..21.146,297174.1..21.805)

```

```

/gene="GLC1A"
/note="first coding exon"
/label=start
join(462..1065,297177.1:21..146,297174.1:21..805)
/gene="GLC1A"
/codon_start=1
/label=CDSGLA1A
/protein_id="CAB09899.1"
/db_xref="GI:2425157"
/db_xref="SWISS-PROT:O99972"
/translacion="MRFECARCCSFCEGPMAYOLLACLAVDYGATAOILRANODS
GRCOYTSVASPNESSCPEBOSQASVTHNLRDSDTORLDEAKNASLESLEHLOU
TLDOAPARPOETQLORELRGLTRREROLLETQRELAVALNLRNLSLAEKELR
QENLNARLRLESSQEVARLRNRGCOPTQRLARAVPGSRVSNWINDTALAEKKE
LTPYAPGRIKESQSGYLRSGCEDTCGCELVWVMEPTLTLTAETITKCYVMWDDPP
LYEPTQETWRIIDVQGVDFVROFEYDILISGFMQEPKRLHLPLESTGAVVYSGSL
YFOGASERTVIRYLENTETVYKAEKIEGAGHGQEPFSGMGYTDIDLAVDAGLWTV
TSDPARKAIVLSKLNPEMLEEOTWENINRQSOANAFTICGTLTVTSYSDATATVN
FADYDGCISIKTLTIPTPKNRKYSMDYNDPLEKELAMNDLNMVTDILNSKA"

```

Query Match	7.7%	Score	407.8	DB	66	Length	1086
Best Local Similarity	98.6%	Pred	No	1.7e-79			
Matches	423	Conservative	0	Mismatches	2	Indels	4
						Gaps	1

[illegible]

RESULT	7
AC007688/c	
LOCUS	
DEFINITION	AC007688 161577 bp DNA PRI 01-SEP-1999
ACCESION	Homo sapiens lp12-27.2-31.7 BAC RPC111-392P7 (Roswell Park Cancer Institute Human BAC library) complete sequence.
VERSION	AC007688
KEYWORDS	AC007688.15 GI:5815499
SOURCE	HTG.
ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. I (bases 1 to 161577)
REFERENCE	

AUTHORS	TITLE	REFERENCE	AUTHORS	TITLE	REFERENCE	AUTHORS	TITLE	REFERENCE	COMMENT
---------	-------	-----------	---------	-------	-----------	---------	-------	-----------	---------

**AUTHORS**  
Muzny, D., Aronson, A.B., Bouck, J., Bunac, C., Chen, Z., Ding, Y.,  
Dugan, S., Durbin, J., Forcum, J., Garcia, C., Gorrill, J.H.,  
Gorrill, L.L., Hernandez, J., Issar, A., Jackson, L., Kneitz, S.,  
Kondjevski, N., Lau, S., Leal, B., Lee, E., Lichtarge, O., Liu, W.,  
Logan, O., Lu, J., Marondel, I., Martinez, C., Merscher, S., Miller, A.,  
Montgomery, K., Oswal, G., Pampall, L.R., Parish, B. J., Perez, L.,  
Rashid, N.D., Rives, C., Scherer, S.E., Shen, H., Shm, C., Simon, M.,  
Vo, Q., Williamson, A., Morley, K.C., Xhang, A.M., Yang, R., Yu, W.,  
Zhou, X., Kucherlapati, R., Nelson, D. and Gibbs, R.A.

**TITLE**  
Direct Submision

**JOURNAL**  
Unpublished

**REFERENCE**  
2. (bases 1 to 161577)

**AUTHORS**  
Morley, K.C.

**JOURNAL**  
Direct Submision

**TITLE**  
Submitted (01-JUN-1999) Molecular and Human Genetics, Baylor  
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3. (bases 1 to 161577)

**REFERENCE**  
Morley, K.C.

**AUTHORS**  
Direct Submision

**JOURNAL**  
Submitted (01-SEP-1999) Human Genome Sequencing Center, Department

**COMMENT**

**CLONE LENGTH:** This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

### Features Listing.

## Features Listing.

### Features Listing.

### Features Listing.

### ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

local mapping efforts.

local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

unpublished.) for Human and Mouse sequences

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect  $< 1e-34$ ) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

Identical matches are annotated as similar.

Identical matches are annotated as similar.

Identical matches are annotated as similar.

Identical matches are annotated as similar.

Identical matches are annotated as similar.

**SEQUENCING-READ COVERAGE:** Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

annotation as Low Coverage.

annotation as Low Coverage.

annotation as Low Coverage.

annotation as Low Coverage.

annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES. This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

html.

html.

html.

html.

QUALSTAT-REPORT-----

```

----- Summary Statistics -----
Contig length:                161577
Phrap values in estimate:     160751
Average error rate (BCR-Phrap estimate): 0.000163681
Fraction of Phrap values less than 40 : 0.0376047
Number of consensus changing edits: 30
Number of N's in consensus : 0

```

Number of N's in consensus : 0

Number of N's in consensus : 0

Number of N's in consensus : 0

Number of N's in consensus : 0

Number of N's in consensus : 0

----- Consensus changing edits -----



Position	Original+Context	Edited+Context
7033	acctgcgctt(n)ccgcccctt	acctgcgctt(c)ccgcccctt
47367	aaaaaaaaa(n)ggaaataat	aaaaaaaaa(a)ggaaataat
51135	aaagaaagaa(n)aaagaaagaa	aaagaaagaa(a)aaagaaagaa
75582	aaaaaaaaa(n)aaagatgtt	aaaaaaaaa(a)aaagatgtt
75585	aaaaaaaaa(n)ggatgttcgc	aaaaaaaaa(a)ggatgttcgc
75667	ctaaagcaga(n)taagatctta	ctaaagcaga(a)taagatctta
75752	tttaataatg(g)gcttgctt	tttaataatg(t)gcttgctt
84017	ggagggaaag(n)aaagaaagag	ggagggaaag(g)aaagaaagag
85227	tttgtttt(n)tgctttttt	tttgtttt(t)tgctttttt
96981	ggagcttcaag(n)atcaatgtag	ggagcttcaag(g)atcaatgtag
111307	aatcctctat(n)cccgaattca	aatcctctat(g)cccgaattca
111374	accagatcac(n)acctttttt	accagatcac(t)acctttttt
111378	gtatccatc(n)ttttttttt	gtatccatc(c)ttttttttt
112621	aaaaaaaaa(n)ccactctaga	aaaaaaaaa(a)ccactctaga
135821	atccaccct(c)ttttttttt	atccaccct(c)ttttttttt
137207	ltgcgaagc(n)cgccaccacag	ltgcgaagc(a)cgccaccacag
137218	cgccaccacag(n)ctgtcttaagt	cgccaccacag(c)ctgtcttaagt
145113	cgaaagtgaa(n)nagatctaat	cgaaagtgaa(a)nagatctaat
145114	cgagatgaan(n)atgaataa	cgagatgaan(a)atgaataa
145212	ggcacccggt(n)ntcacncag	ggcacccggt(c)ntcacncag
145223	ggaccggt(n)lccancagtl	ggaccggt(c)lccancagtl
145238	gtatcgttcac(n)ccagatcatc	gtatcgttcac(g)ccagatcatc
145332	gcaacatgtt(c)nnccacatc	gcaacatgtt(g)nnccacatc
145337	caacatggtt(n)nnccacatc	caacatggtt(a)nnccacatc
145363	aatatgtn(n)nnccacatc	aatatgtn(a)nnccacatc
145377	gttccacag(n)gtgttcgct	gttccacag(t)gtgttcgct
145387	gtggcgttga(c)ttgtatccct	gtggcgttga(g)ttgtatccct
145505	ttgtctcttg(n)ggagggagat	ttgtctcttg(g)ggagggagat
	aaaaaaaaa(n)gaaaaaaaaa	aaaaaaaaa(a)gaaaaaaaaa

Distribution of Quality < 40 Bases -----

[illegible]

Version: 1.01 qxf0

1.01 qxf0.

FEATURES	Location/Qualifiers
source	1. .161577

```
repeat_region complement(3, .65)
repeat_region /rpt_family="MER4B"
repeat_region 161, .471
repeat_region /rpt_family="ALUSx"
repeat_region complement(.607, .702)
repeat_region /rpt_family="MIR"
repeat_region 703, .838
repeat_region /rpt_family="FLAM_C"
repeat_region complement(.639, .916)
repeat_region /rpt_family="MIR"
repeat_region complement(1205, 1377))
repeat_region /rpt_family="MER104"
STS 1388, .1525
/standard_name="WIAF-75"
```

```

repeat_region 1683..1971 /db_xref="dbSTS:65358"
/rpt_family="AluJo"
repeat_region 1974..2261 /rpt_family="AluJo"
repeat_region 2486..2593 /rpt_family="Aluub"
misc_feature 2610..3214 /rpt_family="MIR"
/note="Region: Unigene cluster similar to AA056332 and
AI074576."
STS 2712..2823 /standard_name="SHGC-44583"
/db_xref="dbSTS:48451"
3185..3491 /rpt_family="Alusq"
gene 4509..5919 /gene="Human 5-hydroxytryptamine7 receptor isoform b mRNA,
U68487."
repeat_region complement(6397..6537)
/rpt_family="L2"
6567..7158 /rpt_family="(CCCC)n"
7094..7240 /rpt_family="(CCG)n"
/rpt_family="join(7241..7357,17979..18117,20062..20242,23659..23933)
gene /gene="Unigene cluster containing N42229, R03325, and
A138069."
repeat_region complement(7770..8250)
/rpt_family="MER44B"
8831..9073 /rpt_family="MER102"
repeat_region complement(9075..9303)
/rpt_family="L2"
9671..10135 /rpt_family="LIR33A"
repeat_region 10426..10592 /rpt_family="MIR"
repeat_region 10625..10689 /rpt_family="MIR"
11094..11322 /rpt_family="MIR"
repeat_region complement(11985..12403)
/rpt_family="L2"

```

Query Match	3.6%	Score 189.8;	DB 10;	Length 161577;
Best Local Similarity	78.5%;	Pred. No. 1.3e-31;		
Matches 227;	Conservative 0;	Mismatches 62;	Indels 0;	Gaps 0;

QY	1280	agggtgaaggctctgtgtcttaaacctactgltatgctctacccgtagctcaactgaacc	1339
Db	29160	AGAGTCTCACTATTATACCCAGGCTGGAAGTGCATGGCGCAATCTGGGCTCATTTGCAMCC	29101
QY	1340	tcgtacctccagggttccaagaacattctctgtctcagaacctccgcgtagctcgtggaactacag	1399
Db	29100	TCGTGCTCCCAAGTTCAMGTGATGTCCTCGCTTAGCTTCCGAGTAGCGGACTACAG	29041
QY	1400	ggcgcaagcccgagcttaattttgttatgtttagtagaagttgggtttccaatattagcccg	1459
Db	29040	GCACACGCCCACTAATTTTGTATTATTTATACAGATGGGGTTTTCGGTATGGCCAG	28991
QY	1460	gctcgtcttgaacctctctgacctcagtgatctacacccaccctccctcaagttgcttgg	1519
Db	28980	GCTGGCTCAAACTCCTGACCTCAGAGTGATCACCCGCCCTCACCTTCCTAAATGCTGTGG	28931
QY	1520	attaaagcatagatgacacagcgccggcccggaaggttcagttctttaataag	1568
Db	28920	ATTACAGGGTGTAGCCACCCACGCCCGGCGCTTGAAAGGGTTTAAAGTAG	28872

RESULT	8		
AC023790/c			
LOCUS	AC023790	193123 bp	DNA
DEFINITION	Hom sapiens chromosome 12 clone RP11-377D9, WORKING DRAFT		25-JUL-2000

SEQUENCE, 35 unordered pieces.  
AC023790  
VERSION AC023790.16 GI:9438256  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 193123)  
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.  
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;  
Muzny, D.M., Adams, C., Bailey, M., Barbara, J., Blankenburg, K.,  
Bodota, B., Bouck, J., Bowie, S., Brooks, A., Buhay, C., Bunc, C.,  
Burkett, C., Burrows, J., Carter, M., Chacko, J., Chen, Z., Cox, C.,  
David, R., Delgado, O., Deshazo, D., Ding, Y., Domah-Rashid, N.,  
Dugan-Rocha, S., Durbin, K.J., Fernandez, C., Ferraruto, D.,  
Furum-Tansey, J., Frantz, P., Ganesh, R., Gorrell, J.H., Gorrell, L.L.,  
Guevara, W., Harris, K., Hernandez, J., Hodgson, A., Hogues, M.,  
Hollway, C., Hosak, H., Jackson, L.E., Jackson, L., Jia, Y., Jones, M.,  
Kelly, S., Kondejewski, N., Kong, Y., Kovar, C., Leal, B., Li, Z.,  
Lichter, O., Liu, J., Liu, W., Logan, O., Lozano, R.J., Lu, J.,  
Lucier, R., Martin, R., Martinez, C., McLeod, M.P., Mel, G., Mel, J.,  
Morris, S., Nash, S., Nelson, A., Nguyen, R., Nguyen, N., Nguyen, S.,  
Oswal, G., Parish, B., Paxton, S., Payton, B., Perez, L., Pu, L.L.,  
Quiles, M., Reiter, D., Rives, M., Samuel, S., Say, J., Scherer, S.,  
Shah, E., Shen, H., Simon, M., Sparks, A., Stamps, A., Sugeng, R.,  
Taber, P., Taylor, T., Vasquez, L., Vinson, R., Vo, O., Waldbach, M.,  
Wellington, S., Weinstein, G., Weinstein, I.R., Williamson, A.,  
Worley, K., Wren, J., Wrensford, G., Yu, W., Zhou, X., Nelson, D. and  
Gibbs, R.  
Direct Submission  
Unpublished  
2 (bases 1 to 193123)  
Worley, K.C.  
Direct Submission  
Submitted (18-FEB-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 25, 2000 this sequence version replaced gi:9255941.  
----- Genome Center -----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information -----  
Center project name: HABM  
Center clone name: RP11-377D9  
----- Summary Statistics -----  
Assembly program: Phrap; version 0.990329  
Consensus quality: 167304 bases at least Q40  
Consensus quality: 176244 bases at least Q30  
Consensus quality: 181341 bases at least Q20  
Estimated insert size: 180050; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-gel estimation  
Quality coverage: 3.1x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 35 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1  
\* 21094: contig of 21094 bp in length  
\* 21095: gap of unknown length  
\* 21195: contig of 12736 bp in length  
\* 33931: gap of unknown length  
\* 34031: contig of 14288 bp in length  
\* 48319: gap of unknown length  
\* 48419: contig of 10381 bp in length  
\* 58899: gap of unknown length  
\* 58900: contig of 9824 bp in length

68724 68823: gap of unknown length  
\* 68824 76073: contig of 7250 bp in length  
\* 76074 76173: gap of unknown length  
\* 76174 82868: contig of 6695 bp in length  
\* 82869 82968: gap of unknown length  
\* 82969 90046: contig of 7078 bp in length  
\* 90047 90147: gap of unknown length  
\* 90148 96142: contig of 5995 bp in length  
\* 96143 96241: gap of unknown length  
\* 96242 101873: contig of 5632 bp in length  
\* 101874 101973: gap of unknown length  
\* 101974 108002: contig of 6029 bp in length  
\* 108003 108102: gap of unknown length  
\* 108103 112903: contig of 4801 bp in length  
\* 112904 113004: gap of unknown length  
\* 113005 120209: contig of 7206 bp in length  
\* 120210 120309: gap of unknown length  
\* 120310 126353: contig of 6044 bp in length  
\* 126354 126453: gap of unknown length  
\* 126454 132888: contig of 6435 bp in length  
\* 132889 132988: gap of unknown length  
\* 132989 140137: contig of 7149 bp in length  
\* 140138 140237: gap of unknown length  
\* 140238 144632: contig of 4394 bp in length  
\* 144633 144731: gap of unknown length  
\* 144732 150312: contig of 5581 bp in length  
\* 150313 150412: gap of unknown length  
\* 150413 154114: contig of 3702 bp in length  
\* 154115 154214: gap of unknown length  
\* 154215 158458: contig of 4244 bp in length  
\* 158459 158558: gap of unknown length  
\* 158559 161385: contig of 2827 bp in length  
\* 161386 164320: gap of unknown length  
\* 164321 164420: contig of 2835 bp in length  
\* 164421 167714: gap of unknown length  
\* 167715 167815: contig of 3294 bp in length  
\* 167816 171014: gap of unknown length  
\* 171015 171114: contig of 3200 bp in length  
\* 171115 173843: gap of unknown length  
\* 173844 173943: contig of 2729 bp in length  
\* 173944 176632: gap of unknown length  
\* 176633 176731: contig of 2688 bp in length  
\* 176732 178648: gap of unknown length  
\* 178649 178748: contig of 1917 bp in length  
\* 178749 180232: gap of unknown length  
\* 180233 180332: contig of 1484 bp in length  
\* 180333 183104: gap of unknown length  
\* 183105 183204: contig of 2772 bp in length  
\* 183205 184720: gap of unknown length  
\* 184721 184820: contig of 1516 bp in length  
\* 184821 186137: gap of unknown length  
\* 186138 186237: contig of 1317 bp in length  
\* 186238 186238: gap of unknown length  
\* 186239 188213: contig of 1976 bp in length  
\* 188214 188313: gap of unknown length  
\* 188314 190073: contig of 1760 bp in length  
\* 190074 190173: gap of unknown length  
\* 190174 191855: contig of 1682 bp in length  
\* 191856 191955: gap of unknown length  
\* 191956 193123: contig of 1168 bp in length.  
Location/Qualifiers  
1. 193123  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-377D9"

BASE COUNT 54294 a 41007 c 41291 g 53110 t 3421 others  
ORIGIN  
Query Match 3.6% Score 189.8; DB 77; Length 193123;  
Best Local Similarity 78.5% Pred. No. 1.3e-31;  
Matches 227; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

OY	1280	aggfagagagcctgtgtgtcttaccactacactcgttagctcttaacactgagcctacatgaacc	1339
Db	134658	AGAGCTCCTACCTCTATTATTCACCGCTAGAGTGCATGTGGCCGCAATCTCGCTCATATTGCAACC	1346539
OY	1340	ctctgctccccaaggttcaagcaatctctctgtctcgaagctccgcgcgtatgctggactacag	1399
Db	134638	TCTGCTCCCGAGGTTCAGATGATTCTCTCTGCTTACGTTTCCCGAGTACGTGGGACTACAG	134579
OY	1400	gcgcagccgcggcctaattttgtatctgttagtagaagatgggggtttccacatattagccg	1459
Db	134578	GCACACGCCCGACGTAAATTTTGTATTTTTAACTACAGATGGGGTTCCTCCATATGGCCAG	134519
OY	1460	gctggtcttgaaacctccctactcaagtgatattccaccaccccttcagctcctaagtgctg	1519
Db	134518	GCTGCTCTCAANACTTCCTTACCTCAGGTGATCACCACCCGCTTCAGCTCTTAAATGCTGGG	134459
OY	1530	attcaaggcatgaagltcacacgcgcgcgcgcgaagggttcagtgcttaataag	1568
Db	134458	ATTACAGGTGTGAGACACACCGCCGGCTGTGAAGAAGGTTTAAAGTAG	134410
RESULT	9		
AC012404			
LOCUS	AC012404	DNA	HTG
DEFINITION	AC012404	Homo sapiens chromosome 15 clone RP11-39M21 map 15, WORKING DRAFT	26-MAY-2000
ACCESSION	AC012404	SEQUENCE, 6 unordered pieces.	
VERSION	AC012404.4	GI:7107808	
KEYWORDS	HTG: HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.		
AUTHORS	Bitnen,B., Linton,L., Nusbaum,C. and Lander,E.		
JOURNAL	Homo sapiens chromosome 15, clone RP11-39M21		
REFERENCE	2 (bases 1 to 199722)		
AUTHORS	Bitnen,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeRellano,K., Dewar,K., Domingo,M., Doneelan,L., Doyle,M., Ferreira,P., FitzHugh,M., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heathford,A., Horton,L., Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatsas,A., Klein,J., Lebecky,J., Lien,C., Locke,K., MacDonald,P., Margulis,N., McMan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldin,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Teefaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-Oct-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
COMMENT	On Feb 28, 2000 this sequence version replaced g1:6514012. All repeats were identified using RepeatMasker: Smit, A.P.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html		
	Genome Center		
	Center: Whitehead Institute/ MIT Center for Genome Research		
	Center code: MIBR		
	Web site: http://www-seq.wi.mit.edu		
	Contact: sequence.submissions@genome.wi.mit.edu		
	Project Information		
	Center project name: L261		
	Center clone name: 39_M_21		
	Summary Statistics		
	Sequencing vector: M13: M7815: 100% of reads		
	Chemistry: Dye-terminator Big Dye, 100% of reads		
	Assembly program: Phrap, 0.960731		
	Consensus quality: 178894 bases at least Q40		

[illegible]

Oy 1530 tgaatcaccgcgcgcgaaggctcagtgttaata 1566  
 1111 1111 1111 1111 1111 1111 1111 1111  
 Db 184859 TGAGCCACCATGCCAGCCCATCTTACTTTTACA 184895

RESULT 10  
 AC004973/c 97037 bp DNA PRI 21-DEC-1999  
 LOCUS AC004973/c  
 DEFINITION Homo sapiens PAC clone RP5-113911 from Xq23, complete sequence.  
 AC004973.1 GI:3694660  
 VERSION AC004973.1  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 97037)  
 AUTHORS Sulston, J.E. and Waterston, R.  
 JOURNAL Toward a complete human genome sequence  
 Genome Res. 8 (11), 1097-1108 (1998)  
 REFERENCE 2 (bases 1 to 97037)  
 MEDLINE 99063792  
 AUTHORS Courtney, L., Langston, Y. and Drone, K.  
 JOURNAL The sequence of Homo sapiens PAC clone RP5-113911  
 Unpublished  
 REFERENCE 3 (bases 1 to 97037)  
 AUTHORS Waterston, R.H.  
 JOURNAL Direct Submission  
 Submitted (12-JUN-1998) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 REFERENCE 4 (bases 1 to 97037)  
 AUTHORS Waterston, R.  
 JOURNAL Direct Submission  
 Submitted (03-OCT-1998) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 REFERENCE 5 (bases 1 to 97037)  
 AUTHORS Waterston, R.  
 JOURNAL Direct Submission  
 Submitted (18-MAR-1999) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 REFERENCE 6 (bases 1 to 97037)  
 AUTHORS Waterston, R.  
 JOURNAL Direct Submission  
 Submitted (21-DEC-1999) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Oct 3, 1998 this sequence version replaced gi:3213024.  
 COMMENT  
 -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc  
 Contact: saplens@wustl.wustl.edu  
 -----  
 Summary Statistics  
 -----  
 Center project name: H\_DJ1139101

NOTICE: This sequence may not represent the entire insert of this  
 clone. It may be shorter because we only sequence overlapping  
 clone sections once, or longer because we provide a small overlap  
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.

MAPPING INFORMATION:  
 This sequence was generated from part of bacterial clone contigs of  
 human chromosome X, constructed by the chromosome X mapping group

at the Sanger Centre, Wellcome Trust Genome Campus, Hinxton, UK.  
 Further information can be found at  
<http://www.sanger.ac.uk/HGP/ChrX/>

#### SOURCE INFORMATION:

This clone was derived from human PAC library RPC1-5, prepared by  
 Pieter de Jong and coworkers at the Roswell Park Cancer Institute  
 (http://bacpac.med.buffalo.edu) using the method described by  
 Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from  
 one male donor.  
 The clone may be obtained either from Genome Systems, Inc.  
 (http://www.genomesystems.com) or Research Genetics, Inc.  
 (http://www.resgen.com); or from Pieter de Jong.  
 VECTOR: pCYPAC2

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP4-555N2; the clone sequenced  
 to the right is RP3-404F18, 200 bp overlap. Actual start of this  
 clone is at base position 1 of RP5-113911; actual end is at 13269  
 of RP3-404F18.

#### FEATURES

source	Location/Qualifiers
1. 97037	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="X"
	/map="Xq23"
	/clone="RP5-113911"
	/clone.lib="RPC1-5"
repeat_region	530..801
	/rpt_family="L2"
repeat_region	830..1035
	/rpt_family="Alu"
repeat_region	2179..2349
	/rpt_family="MER1_type?"
repeat_region	3434..3489
	/rpt_family="MIR"
repeat_region	3526..3739
	/rpt_family="MIR"
repeat_region	4681..4852
	/rpt_family="MIR"
repeat_region	5395..5961
	/rpt_family="MER4-group"
repeat_region	6293..6393
	/rpt_family="L1"
repeat_region	6394..7105
	/rpt_family="L1"
repeat_region	7104..7211
	/rpt_family="L1"
repeat_region	7342..7520
	/rpt_family="L1"
repeat_region	7542..7644
	/rpt_family="Alu"
repeat_region	7645..8674
	/rpt_family="Retroviral"
repeat_region	8682..9314
	/rpt_family="Retroviral"
repeat_region	9315..9368
	/rpt_family="7SLRNA"
repeat_region	9371..10137
	/rpt_family="Retroviral"
repeat_region	10165..10817
	/rpt_family="Retroviral"
repeat_region	10818..10973
	/rpt_family="Alu"
repeat_region	10974..11111
	/rpt_family="L1"
repeat_region	11238..11612
	/rpt_family="MALR"
repeat_region	11972..12283
	/rpt_family="Alu"
repeat_region	12570..12854
	/rpt_family="L1"
repeat_region	12859..13433
	/rpt_family="MER1_type"

```
repeat_region 13435..13564
/rpf_family="Alu"
repeat_region 13580..13845
/rpf_family="MERL_type"
repeat_region 13850..14118
/rpf_family="L1"
repeat_region 14119..14411
/rpf_family="Alu"
repeat_region 14412..14813
/rpf_family="L1"
repeat_region 14801..15447
/rpf_family="L1"
repeat_region 16501..16654
/rpf_family="L1"
repeat_region 17412..17533
/rpf_family="L2"
repeat_region 17534..17793
/rpf_family="MERL_type"
repeat_region 17794..17912
/rpf_family="Alu"
repeat_region 17933..18249
/rpf_family="MERL_type"
repeat_region 18293..18358
/rpf_family="L1"
repeat_region 18348..18476
/rpf_family="MIR"
repeat_region 18483..18908
/rpf_family="L2"
repeat_region 18932..19227
/rpf_family="MALR"
repeat_region 19390..19604
/rpf_family="Alu"
repeat_region 19606..20227
/rpf_family="L2"
repeat_region 20228..20477
/rpf_family="Achobo"
repeat_region 20478..20721
/rpf_family="Alu"
repeat_region 20738..20828
/rpf_family="Achobo"
repeat_region 20848..21707
/rpf_family="Achobo"
repeat_region 21717..22111
/rpf_family="L2"
repeat_region 22163..22378
/rpf_family="L2"
repeat_region 22866..23031
/rpf_family="L2"
repeat_region 23060..23339
/rpf_family="L1"
repeat_region 23409..24201
/rpf_family="Alu"
repeat_region 24203..24554
/rpf_family="L2"
repeat_region 24572..25105
/rpf_family="Achobo"
repeat_region 25106..25479
/rpf_family="L1"
repeat_region 25480..25768
/rpf_family="Alu"
repeat_region 25769..25985
/rpf_family="L1"
/rpf_family="L1"
```

Query Match 3.58: Score 187.8; DB 9; Length 97037;  
Best Local Similarity 78.78; Pred. No. 3.4e-31;  
Matches 240; Conservative 0; Mismatches 57; Indels 8; Gaps 1;

```
Oy 1280 aggggaaggtcgtgtcttaccctaccctacgtatgctctacacacgtgagctcagctcaaac 1339
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54922 AGAGCTTCTGCTGTCAACGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGC 54863
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1340 tctgcctccaggttcaggaattctctgtctcagcctccgcgtagctgggactacag 1399
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 54862 TCTGCTCTCTGGGTTACAGCAATTCTCCTGTCTGACGCTCCGAGTAGCTGGACTACAG 54803
Oy 1400 gcg-----cagccgcgcgttaatttctgattgtagtagagatgggtttaccata 1451
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54802 GCGCATGCGACACGACCCAGCTAATTTTGTATTTTATTTAGTATGAGATGGGTTTACCATA 54743
Oy 1452 ttaagccgcgtgtcttgaactctgacctcagtgatgataccaccactcagcctctaa 1511
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54742 TTGCTCAGGCTGTCTGAACTCCTGACCTCAGATGATCCACCTGCTCGCCCTCCCAA 54683
Oy 1512 gtgtctggaattacagcgatgtagtcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1571
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54682 GTGCTGGATTACAGGCGTGCAGCCGCGCCGCTCTTCTAGTATTTTGGAAATGTA 54623
Oy 1572 taact 1576
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54622 CAATT 54618
```

RESULT 11  
HUMYWXD703/c  
LOCUS HUMYWXD703 135038 bp DNA PRI 24-DEC-1996  
DEFINITION Homo sapiens ADP/ATP carrier protein (ANT-2) gene, complete cds.  
ACCESSION L/8810  
VERSION L/8810.1 GI:1381111  
KEYWORDS  
SOURCE  
ORGANISM  
human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 135038)  
Chen,C.N., Su,Y., Baydayan,P., Siruno,A., Nagaraja,R.,  
Mazarella,R., Schlessinger,D. and Chen,E.  
Ordered shotgun sequencing of a 135 kb Xq25 YAC containing ANT2 and  
four possible genes, including three confirmed by EST matches  
Nucleic Acids Res. 24 (20), 4034-4041 (1996)

JOURNAL  
MEDLINE  
COMMENT  
Submitted by:  
Ellison Chen,  
Advanced Center for Genetic Technology,  
Applied Biosystems Division of Perkin Elmer Corp., 850 Lincoln  
Center Drive,  
Foster City, CA 94404 USA  
and  
David Schlessinger,  
Department of Molecular Microbiology and Center for Genetics in  
Medicine  
Washington University School of Medicine,  
St. Louis MO 63110 USA  
e-mail: elison@geneseq.apidbio.com and david@genetics.wustl.edu  
Note: Gene predictions were accomplished with runs of Grail  
versions 1.1 and 1.2, coupled with fasta and blastx comparisons to  
genbank & non-redundant peptide libraries. Repeat analysis was  
accomplished via censor.  
The Rat EST105369 shows significant homology via blastx to this  
sequence

Strand Start End  
top 9072 9326  
The Graves Disease carrier protein (X66035) shows significant  
homology via blastx to this sequence  
Strand Start End  
top 28802 29076  
top 35858 36081  
Comments for gene ANT-2:  
This gene shows homology via blastx to the EST clone y115a12 Strand  
Start End  
bottom 125256 125915.

FEATURES  
source  
repeat\_unit  
complement(271..560)

repeat_unit	/rpt_family="Alu-Sx" /evidence-experimental complement(774. .1726) /rpt_family="L1" /evidence-experimental complement(1842. .2131) /rpt_family="L1" /evidence-experimental 2312. .2465	repeat_unit	/evidence-experimental complement(15805. .16906) /rpt_family="LTRL2" /evidence-experimental 17588. .17930 /rpt_family="MER44" /evidence-experimental complement(20000. .20276) /rpt_family="Alu-Sb0 or an Alu-Sb1" /evidence-experimental 20353. .20856
repeat_unit	/rpt_family="L1PA11" /evidence-experimental complement(2468. .2757) /rpt_family="Alu-Sx" /evidence-experimental 2774. .3059	repeat_unit	/rpt_family="LIME3A" /evidence-experimental 20900. .20986 /rpt_family="MSTA" /evidence-experimental 20988. .21276
repeat_unit	/rpt_family="MUTIC" /evidence-experimental complement(3204. .3325) /rpt_family="MIR2" /evidence-experimental 3384. .3674	repeat_unit	/rpt_family="Alu-Sx" /evidence-experimental 21342. .21723 /rpt_family="LIME3A" /evidence-experimental 21841. .22129
repeat_unit	/rpt_family="Alu-Sg" /evidence-experimental 3794. .4216 /rpt_family="LIMA9" /evidence-experimental 7502. .7792	repeat_unit	/rpt_family="Alu-Sp" /evidence-experimental 22150. .22445 /rpt_family="Alu-U" /evidence-experimental complement(22542. .22789)
repeat_unit	/rpt_family="Alu-U" /evidence-experimental 7965. .8072 /rpt_family="LIMA2" /evidence-experimental 8073. .8360	repeat_unit	/rpt_family="THE1B" /evidence-experimental complement(23235. .23377) /rpt_family="MIR2" /evidence-experimental complement(23513. .23798)
repeat_unit	/rpt_family="Alu-Sx" /evidence-experimental 8379. .8535 /rpt_family="LIMA5" /evidence-experimental 8540. .8743	repeat_unit	/rpt_family="Alu-Sx" /evidence-experimental 24296. .24578 /rpt_family="Alu-Sx" /evidence-experimental 24808. .24944
repeat_unit	/rpt_family="L1" /evidence-experimental 8744. .8869 /rpt_family="Alu-U" /evidence-experimental complement(9771. .10055)	repeat_unit	/rpt_family="Alu-Sg" /evidence-experimental 24945. .25234 /rpt_family="Alu-Sb2" /evidence-experimental 25236. .25393
repeat_unit	/rpt_family="Alu-Sx" /evidence-experimental 10292. .10780 /rpt_family="L1" /evidence-experimental 11403. .11692	repeat_unit	/rpt_family="Alu-Sc" /evidence-experimental complement(26486. .26773) /rpt_family="Alu-Sx" /evidence-experimental 26570. .26859
repeat_unit	/rpt_family="Alu-Sx" /evidence-experimental 12097. .12459 /rpt_family="MSR" /evidence-experimental complement(12492. .12701)	repeat_region	/rpt_family="CPG Island" /evidence-experimental 26801. .27091 /rpt_family="Alu-Sx" /evidence-experimental 27866. .27986
repeat_unit	/rpt_family="Alu-U or an Alu-S" /evidence-experimental complement(12730. .13020) /rpt_family="Alu-Sx" /evidence-experimental 13708. .13997	repeat_unit	/rpt_family="MIR" /evidence-experimental complement(28015. .28185) /rpt_family="MIR" /evidence-experimental 28550. .29215
repeat_unit	/rpt_family="Alu-Sg" /evidence-experimental complement(14402. .14541) /rpt_family="MIR2" /evidence-experimental 14546. .14836	repeat_region	/rpt_family="CPG Island" /evidence-experimental 31026. .31323 /rpt_family="MIR" /evidence-experimental complement(32698. .32984)
repeat_unit	/rpt_family="Alu-Sx" /evidence-experimental 15348. .15628 /rpt_family="Alu-Sx"	repeat_unit	/rpt_family="Alu-Sc" /evidence-experimental

Repeat_Unit	33130..33399	Query Match	3.5%; Score 187.8; DB 67; Length 135038;
	/rpt_family="MER21"	Best Local Similarity 78.7%; Pred. No. 3.5e-31;	
	Matches 240; Conservative 0; Mismatches 57; Indels 8; Gaps 1;		
Db	1280	agggtgaggggtctgtgtcttacacctactgtatgtcttaaccctgagctactcgtcaacc	1339
Db	8349	AGAGCTTGTGCTTGTTCATCCAAAGCTGGAGATGCAAGTGCACGATCTTCAGTCACTCCAAAC	8290
Db	1340	tctgtctccaggttcacaaactctctcgtctcagccctcccgctgaactgagctacag	1399
Db	8289	TCTGCTCTCTGGGCTTCACCAATTCTCTCTCAACCTCCCGAGTAGCTGGGACTACAG	8230
Db	1400	gcgc-----cagcccgccctaattttgtatctgttagtagagatggggtttaccata	1451
Db	8229	GGCATATGCCACACGCCCGCTAATTTTGTATTATTAGATAGATGGGGTTTCACATA	8170
Db	1452	ttagccgagctgtgtcttgacctccgacctgaaggatcaccacactagctctctaa	1511
Db	8169	TTGGTCAGGCTGTGTTGAATCCCTGACTCCAGATGATCCACTCTCCCTGCCCTCCAA	8110
Db	1512	gtgcggaggtacagcagcagtgacccggccgcccaggggtcagtggttaataagaa	1571
Db	8109	GTGCTGGGATTTACAGCGCTGACCGACCGCGCCGCTCTTCTAGTATTTTGGAAATGA	8050
Db	1572	tact 1576	
Db	8049	CATT 8045	
RESULT 12			
LOCUS	HS821D11/c		
DEFINITION	Human DNA sequence from clone RP5-821D11 on chromosome 22q12.3-13.1. Locus contains three partial unknown genes, one downstream of a predicted Cpg Island, and the first coding exon of the SREBF2 gene for Sterol Regulatory Element Binding Transcription Factor 2 downstream of a predicted Cpg Island, ESTs, STSs, GSSs and genomic marker D22S157./		
ACCESSION	AL021453		
VERSION	AL021453.1 GI:3413288		
KEYWORDS	HTG; D22S157; SREBF2; Sterol Regulatory Element Binding Transcription Factor 2.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 76727)		
TITLE	Bird,C.		
JOURNAL	Direct Submission		
COMMENT	Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire, CB30 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk On Aug 12, 1998 this sequence version replaced gi:3355590. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome Mapping Group. Further information can be found at <a href="http://www.sanger.ac.uk/HGP/Chr22">http://www.sanger.ac.uk/HGP/Chr22</a> During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession		

Numbers given in the feature table with their source database: Em: EMBL, Sw: SWISSPROT, Tr: TREMBL, Wp: WORMPEP, information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C.elegans/wormpep> RP5-821D11 is from the library RPc1-5 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/> VECTATOR: PCYPAC2

**IMPORTANT:** This sequence is not the entire insert of clone RP5-821D11. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

The true left end of clone RP5-821D11 is at 1 in this sequence. The true left end of clone CTA-350D10 is at 7628 in this sequence. The true right end of clone CTA-10956 is at 42082 in this sequence. The start of this sequence overlaps with sequence 299716 The end of this sequence overlaps with sequence 283840.

Location/Qualifiers

FEATURES	Source
repeat_region	location/Qualifiers 1..76727 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="22" /map="q12.3-13.1" /clone="RP5-821d11" /clone_1bp="RRC1-5" 350..434 /note="L2 repeat: matches 2490..2537 of consensus" 435..743 /note="AluIb repeat: matches 8..312 of consensus" 446..448 /note="clone CTA-109G6 tct in this entry substitution" /replace="ttt" 744..840 /note="L2 repeat: matches 2381..2490 of consensus" join(1230..1711,13695..13736) /gene="dJ821D11.1" /note="match: ESTs: Em:AA419437" /evidence="not_experimental" /product="dJ821D11.1 (PUTATIVE protein)" 1230..13736 /gene="dJ821D11.1" 1314..1316 /gene="dJ821D11.1" /note="clone CTA-109G6 tgt in this entry substitution" /replace="tct" join(1601..1711,13695..13736) /gene="dJ821D11.1" /note="other possible startcodon at 1583 this gene and dJ821D11.2 could be part of one gene match: proteins: Tr:O95505" /codon_start=1 /evidence="not_experimental" /product="dJ821D11.1 (PUTATIVE protein)" /protein_id="CA16279.1" /db_xref="GI:4200330" /db_xref="SPTREMBL:O95505" /translation="MLALTLAKADSPITALCSAMLLTASFSNQHKGSIQKQPLSLQ ACVGCLE" 1760..1762 /gene="dJ821D11.1" /note="clone CTA-109G6 gag in this entry substitution" /replace="ggg" 1773..1923 /note="MIR repeat: matches 47..201 of consensus" 1901..1903 /gene="dJ821D11.1" /note="clone CTA-109G6
variation	
gene	
variation	
CDS	
variation	
repeat_region	
variation	
repeat_region	
variation	



repeat_region	tca in this entry substitution" /replace="tta" 2010. .2209 /note="L2 repeat: matches 2250. .2453 of consensus" 2302. .2304 /gene="dj821D11.1" /note="clone CTA-109G6 9ga in this entry substitution" /replace="gaa" 3042. .3152 /note="L1M1 repeat: matches 6231. .6332 of consensus" 3162. .3457 /note="AluSP repeat: matches 1. .297 of consensus" 3458. .3551 /note="MIR repeat: matches 74. .148 of consensus" 3552. .3848 /note="Aluub repeat: matches 1. .299 of consensus" 3849. .3887 /note="MIR repeat: matches 32. .74 of consensus" 4753. .4754 /gene="dj821D11.1" /note="clone CTA-109G6 tt in this entry deletion" /replace="ttgagt" 4772. .4790 /note="L1M4 repeat: matches 3167. .3185 of consensus" 4791. .5099 /note="AluSG repeat: matches 2. .312 of consensus" 4791. .4792 /gene="dj821D11.1" /note="clone CTA-109G6 tt in this entry deletion" /replace="ttttattat" 5100. .5116 /note="L1M4 repeat: matches 3151. .3167 of consensus" 5107. .5112 /gene="dj821D11.1" /note="clone CTA-109G6 ttttat in this entry insertion" /replace="tt" 5117. .5427 /note="Aluab8 repeat: matches 1. .318 of consensus" 5359. .5361 /gene="dj821D11.1" /note="clone CTA-109G6 att in this entry substitution" /replace="act" 5430. .5729 /note="AluSP repeat: matches 1. .300 of consensus" 5730. .5764 /note="L1M4 repeat: matches 3116. .3151 of consensus" 5761. .5762 /gene="dj821D11.1" /note="clone CTA-109G6 at in this entry deletion" /replace="gtt" 5765. .6075 /note="AluY repeat: matches 1. .311 of consensus" 5772. .5773 /gene="dj821D11.1" /note="clone CTA-109G6 tt in this entry deletion" /replace="tat" 5778. .5779 /gene="dj821D11.1"
---------------	---

```

variation
    /note="clone CTA-109G6
    it in this entry
    deletion"
    /replace="ttaa"
    5780..5782
    /gene="d4821D11.1"
    /note="clone CTA-109G6
    tat in this entry
    substitution"
    /replace="ttt"
    5799..5801
    /gene="d4821D11.1"
    /note="clone CTA-109G6
    cgg in this entry
    substitution"

```

Query Match	3.5%	Score 185.8;	DB 65;	Length 76727;
Best Local Similarity	78.5%;	Pred. No. 9.3e-31;		
Matches 238;	Conservative	0;	Mismatches 57;	Indels 8; Gaps 1

variation	/note="MIR repeat: matches 32. 74 of consensus" 4753. .4754
-----------	--

repeat_region	4772..4790	repeat: matches 3167..3185 of consensus"
repeat_region	/note="11M4	
repeat_region	4791..5099	repeat: matches 2..312 of consensus"
variation	/note="A1ubg	
variation	4791..4792	

Repeat_region	5100. 5116	/note="LM4 repeat: matches 3151. 3167 of consensus"
Variation	5107. 5112	

repeat_region	5117..5427	
/note="AluYb8 repeat: matches 1..318 of consensus"		
variation	5359..5361	
/gene="dJ82D11.1"		
/note="clone CTA-109C6 alt in this entry"		

repeat_region	5430..5729	/note="AluJp repeat: matches 1..300 of consensus"
repeat_region	5730..5764	/note="11M4 repeat: matches 3116..3151 of consensus"
variation	5761..5762	

```

repeat_region      5765..6075
                    /note="Alu repeat: matches 1. .311 of consensus"
variation          5772..5773
                    /gene="J3821D11.1"
                    /note="Clone CRA-10956
                    it in this entry"

```

variation

Qy	1572	taa	1574
Db	57118	AAA	57116

```
Oy 1512 gtcctggaattacagcatgtacccgcgccgccaaggctcagtgtttaataaggaa 1571
      |||||
Db 57178 GTCCTGGGATTACAGGCATGAGCCACTGTGCCCGCCAAATAATCTGGGTAATTTC AATTA 57119
```

QY 1572 taa 1574

RESULT 13  
AC026395/C

LOCUS	AC026395	152044 bp	DNA	HTG	15-JUN-2000
DEFINITION	Homo sapiens chromosome 10 clone RP11-45D20, WORKING DRAFT SEQUENCE, 44 unordered pieces.				

VERSION AC026395.3 CI:8567738  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

ORGANISM    Homo sapiens

REFERENCE  
1 (bases 1 to 15204)

TITLE	Genome The
-------	------------

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 152044)

TITLE	Direct Submission
Journal of Management Education	

COMMENT  
Street, Waltham, MA 02453, USA  
On Jun 15, 2000 this sequence version replaced gi:7330305.

Center: Genome Therapeutics Corporation  
Center code: GTC  
Web site: <http://www.genomecorp.com/>  
Contact: [gtc-seqcenter@genomecorp.com](mailto:gtc-seqcenter@genomecorp.com)

----- Project Information  
Center project name: hg344  
----- Summary Statistics  
Sequencing vector: N/A  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 990315  
Consensus quality: 126801 bases at least Q40  
Consensus quality: 138138 bases at least Q30  
Consensus quality: 140422 bases at least Q20  
Insert size: 147744; sum-of-contigs  
Quality coverage: 3.2x in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 44 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```
1 1006: contig of 1006 bp in length
* 1007 1106: gap of unknown length
* 1107 2475: contig of 1369 bp in length
* 2476 2575: gap of unknown length
* 2576 3927: contig of 1352 bp in length
* 3928 4027: gap of unknown length
* 4028 5333: contig of 1306 bp in length
* 5334 5433: gap of unknown length
* 5434 6664: contig of 1131 bp in length
* 6665 7870: gap of unknown length
* 7871 9384: contig of 1206 bp in length
* 9385 9484: gap of unknown length
* 9485 10791: contig of 1307 bp in length
* 10792 10891: gap of unknown length
* 10892 13010: contig of 2119 bp in length
* 13011 13110: gap of unknown length
* 13111 14617: contig of 1507 bp in length
* 14618 14717: gap of unknown length
* 14718 16217: contig of 1500 bp in length
* 16218 16317: gap of unknown length
* 16318 18106: contig of 1789 bp in length
* 18107 18206: gap of unknown length
* 18207 19498: contig of 1292 bp in length
* 19499 20667: gap of unknown length
* 20668 20967: contig of 1269 bp in length
* 20968 22033: gap of unknown length
* 22034 22303: contig of 1236 bp in length
* 22304 24098: gap of unknown length
* 24099 24198: gap of unknown length
* 24199 26005: contig of 1807 bp in length
* 26006 26105: gap of unknown length
* 26106 27441: contig of 1336 bp in length
* 27442 29037: contig of 1496 bp in length
* 29038 29138: gap of unknown length
* 29139 30660: contig of 1523 bp in length
* 30661 30760: gap of unknown length
* 30761 32755: contig of 1995 bp in length
* 32756 32855: gap of unknown length
* 32856 34483: contig of 1628 bp in length
* 34484 34583: gap of unknown length
* 34584 37578: contig of 2895 bp in length
* 37579 37678: gap of unknown length
* 37679 39699: contig of 2021 bp in length
* 39700 42411: contig of 2612 bp in length
* 42412 42511: gap of unknown length
* 42512 45638: contig of 3127 bp in length
* 45639 45738: gap of unknown length
* 45739 49739: contig of 4001 bp in length
```

```
* 49740 49839: gap of unknown length
* 49840 54276: contig of 4437 bp in length
* 54277 54376: gap of unknown length
* 54377 59051: contig of 4675 bp in length
* 59052 59151: gap of unknown length
* 59152 64626: contig of 5475 bp in length
* 64627 67988: contig of 3262 bp in length
* 67989 68088: gap of unknown length
* 68089 72495: contig of 4406 bp in length
* 72496 72594: gap of unknown length
* 72595 75643: contig of 3049 bp in length
* 75644 75744: gap of unknown length
* 75745 79510: contig of 3767 bp in length
* 79511 79611: gap of unknown length
* 79612 84665: contig of 5054 bp in length
* 84666 84764: gap of unknown length
* 84765 90514: contig of 5750 bp in length
* 90515 90614: gap of unknown length
* 90615 96185: contig of 5571 bp in length
* 96186 96285: gap of unknown length
* 96286 102048: contig of 5763 bp in length
* 102049 102148: gap of unknown length
* 102149 108361: contig of 6213 bp in length
* 108362 108461: gap of unknown length
* 108462 114329: contig of 5868 bp in length
* 114330 114429: gap of unknown length
* 114430 120644: contig of 6215 bp in length
* 120645 120744: gap of unknown length
* 120745 128025: contig of 7281 bp in length
* 128026 128125: gap of unknown length
* 128126 138788: contig of 10663 bp in length
* 138789 138888: gap of unknown length
* 138889 152044: contig of 13156 bp in length.
```

FEATURES  
Source  
1..152044  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/clone="RP11-45D20"  
/clone\_1bp="RPC1-11"  
Location/Qualifiers  
BASE COUNT 40176 a 33928 c 33229 g 40362 t 4349 others  
ORIGIN

Query Match 3.5%: Score 185.8; DB 79; Length 152044;  
Best Local Similarity 80.8%; Pred. No. 9.6e-31;  
Matches 232; Conservative 0; Mismatches 47; Indels 8; Gaps 1;

```
Oy 1321 accctgagctcactgcaaacctctgctcccaaggttcaagcaattctctgtctcagctcc 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30398 ATCTGGGCTACTGCAACCTCCAGCTCCGGGTTCAAGTATCTCCGCTCAGCTCC 30339

Oy 1381 cgcgtagctggaactacagc-----gcagcccgccgaatttctgtcttctgta 1432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30338 TGAGTAGCTGCGATTACAGCCACCCACATACACGCCCAAGCAATTTTGTATTGTA 30279

Oy 1433 gagatggggttccaccatcagcccgctgcttgaacctcgaacctcaggtgataca 1492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30278 GAGATGGGCTTTCACCATGTTGCGCAGCTGCTTGAACCTCGACCTCAGTGATCA 30219

Oy 1493 cccacactcagctcctcaagtctggtgattacaggaatgaatcagccgcccgaag 1552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30218 CCGGCTTAGCGCTCCCAAGCTGGGATTCAGAGCATGACCATGCGACCTGCGCAATT 30159

Oy 1553 gtcaagtcttaataaggaataactgaatggttactcaacaacaag 1599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30158 TTAATTTTTTTAATGAAAAAATGATATTATAAAAAAACACAG 30112
```

RESULT 14  
AC025947/c  
LOCUS AC025947 157057 bp DNA HTG 10-JUN-2000

DEFINITION	Homo sapiens chromosome 10 clone RFL1-78A18, WORKING DRAFT SEQUENCE, 26 unordered pieces.
ACCESSION	AC023947
VERSION	AC023947.3 GI:8439851
KEYWORDS	HTG: HTGS_PHRASE; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 157057) Smith,D.R. Genome Therapeutics Corporation Sequencing Center: Human Genome Sequence Data Unpublished 2 (bases 1 to 157057) Smith,D.R. Direct Submission Submitted (18-MAR-2000) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA On Jun 10, 2000 this sequence version replaced gi:7558340.
JOURNAL	----- Center: Genome Therapeutics Cooperation Reference AUTHORS TITLE JOURNAL
COMMENT	----- Center code: GTC Web site: http://www.genomecorp.com/ Contact: gtc.seqcenter@genomecorp.com ----- Project Information Center project name: hg307 ----- Summary Statistics Sequencing vector: N/A Chemistry: Dye-terminator Big Dye; 100% of reads Assembly program: Phrap; version 990315 Consensus quality: 140586 bases at least Q40 Consensus quality: 148596 bases at least Q30 Consensus quality: 150061 bases at least Q20 Insert size: 154557; sum-of-contigs Quality coverage: 3.7x in Q20 bases; sum-of-contigs  ----- NOTE: This is a 'working draft' sequence. It currently consists of 26 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  1 1544: contig of 1544 bp in length * 1545 1644: gap of unknown length * 1645 3218: contig of 1574 bp in length * 3219 3318: gap of unknown length * 3319 5022: contig of 1704 bp in length * 5023 5122: gap of unknown length * 5123 6553: contig of 1431 bp in length * 6554 6654: gap of unknown length * 6655 8651: contig of 1998 bp in length * 8652 8752: gap of unknown length * 8753 10843: contig of 2092 bp in length * 10844 10944: gap of unknown length * 10945 13772: contig of 2829 bp in length * 13773 13872: gap of unknown length * 13873 15825: contig of 1553 bp in length * 15826 15925: gap of unknown length * 15926 19591: contig of 3666 bp in length * 19592 19691: gap of unknown length * 19692 22350: contig of 2659 bp in length * 22351 22451: gap of unknown length * 22452 25948: contig of 3498 bp in length * 25949 26048: gap of unknown length * 26049 29641: contig of 3593 bp in length * 29642 29741: gap of unknown length * 29742 34105: contig of 4364 bp in length * 34106 34205: gap of unknown length * 34206 38484: contig of 4279 bp in length * 38485 38584: gap of unknown length

	*	38585	41992:	contig of 3408 bp in length
	*	41993	42092:	gap of unknown length
	*	42093	47920:	contig of 5828 bp in length
	*	47921	48020:	gap of unknown length
	*	48021	54416:	contig of 6396 bp in length
	*	54417	54516:	gap of unknown length
	*	54517	60194:	contig of 5678 bp in length
	*	60195	60294:	gap of unknown length
	*	60295	67639:	contig of 7345 bp in length
	*	67640	67739:	gap of unknown length
	*	67740	74790:	contig of 7051 bp in length
	*	74791	74890:	gap of unknown length
	*	74891	85631:	contig of 10741 bp in length
	*	85632	85731:	gap of unknown length
	*	85732	93940:	contig of 8209 bp in length
	*	93941	94040:	gap of unknown length
	*	94041	104191:	contig of 10151 bp in length
	*	104192	104291:	gap of unknown length
	*	104292	116387:	contig of 12096 bp in length
	*	116388	116487:	gap of unknown length
	*	116488	135857:	contig of 19370 bp in length
	*	135858	135957:	gap of unknown length
	*	135958	157057:	contig of 21100 bp in length.
FEATURES				
SOURCE				
			Location/Qualifiers	
			1..157057	
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/chromosome="Chromosome 10"		
		/clone="RP11-78A18"		
		/clone_1bp="RPC1-11"		
BASE COUNT	45178 a	32550 c	32372 g	44448 t 2509 others
ORIGIN				
Query Match		3.5%:	Score 185.8; DB 78;	Length 157057;
Best Local Similarity		80.8%:	Pred. No. 9.7e-31;	
Matches 232;	Conservative	0;	Mismatches 47;	Indels 8; Gaps 1;
QY	1321	acctgagctcactgcaacccctgcctcccagggttcgaagaattctctcgtgtcagcctcc	1380	
Db 109551	AMTCGGCTCACTGCACCTCCACTCACCCTCCCGGAGTTCAGATTCCTCGCTCAGCCTCC	109492		
QY	1381	cgcgttagctgggaactaacagc-----gaacgccggctaatttttgtatgttaagta	1432	
Db 109491	TGAGTAGCTGGGATTACAGGCCACCACACTCACACGCCCGCATTAATTTTGATTTTTTAATA	109432		
QY	1433	ggagtgggggttccaccatatagccccgcctgtgtcttgaaaccccgaaactcagggatcac	1492	
Db 109431	GAGATGGGGTTTCCACCAATGTGGCCAGGCTGGTCTTGAACTCTGACCTCAAGTGATCCA	109372		
QY	1493	cccacactcagcctcccttaaagtgtctggaattcaagcatgaqtacacgcgcgcgccaag	1552	
Db 109371	CCCCGCTTAAGCTCCCAAGAAGTGCTGGATTACAGGCATGACGACCACTGCACCTGGCCATT	109312		
QY	1553	gtcagtgctttaataaggaataacttgaatggtttacttaaccaacaag	1599	
Db 109311	TTAATTTTAAAAATGAAGAAAATGTGGTATTTTAAAAAAAACACAG	109265		
RESULT 15				
ACOL15488				
LOCUS	ACOL15488	161499 bp	DNA	HTG 16-MAR-2000
DEFINITION	Homo sapiens clone RP11-20F6,	WORKING DRAFT SEQUENCE,	16 unordered pieces.	
ACCESSION	ACOL15488			
VERSION	ACOL15488.4 GI:7107977			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 161499)			
AUTHORS	Birtten,B., Jinton,L., Nusbaum,C. and Lander,E.			

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Homo sapiens, clone RP11-20F6  
Unpublished  
2 (bases 1 to 161499)  
Bairden,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckery,R., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Castelle,A., Colangelo,M., Collins,S., Collymore,A.,  
Cooke,P., DeArrellano,K., Dewar,K., Domino,M., Donejan,L., Doyle,M.,  
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,  
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Harford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karkates,A., Klein,J.,  
Lenczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
McMan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,  
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thoman,N., Stojanovic,N., Sudramanah,A., Talamas,J.,  
Testafye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

JOURNAL  
COMMENT

Submitted (16-Nov-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 28, 2000 this sequence version information replaced g1:6715933.  
All repeats were identified using RepeatMasker:  
Smit,A.F.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center -----  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information -----  
Center project name: L3895  
Center clone name: 20\_F\_6  
----- Summary Statistics -----  
Sequencing vector: M13; M7781S; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152763 bases at least Q40  
Consensus quality: 156653 bases at least Q20  
Consensus quality: 158136 bases at least Q20  
Insert size: 160000; agarose-ff  
Insert size: 159999; sum-of-contigs  
Quality coverage: 5.9 in Q20 bases; agarose-ff  
Quality coverage: 5.9 in Q20 bases; sum-of-contigs  
-----  
NOTE: This is a 'working draft' sequence. It currently  
consists of 16 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1	2206:	contig of 2206 bp	in length
*	2207	2306: gap of 100 bp	
*	2307	6910: contig of 4604 bp	in length
*	6911	7010: gap of 100 bp	
*	7011	9818: contig of 2808 bp	in length
*	9819	9918: gap of 100 bp	
*	9919	13394: contig of 3476 bp	in length
*	13395	13494: gap of 100 bp	
*	13495	17122: contig of 3628 bp	in length
*	17123	17222: gap of 100 bp	
*	17223	19752: contig of 2530 bp	in length
*	19753	19852: gap of 100 bp	
*	19853	22958: contig of 4106 bp	in length
*	23959	24058: gap of 100 bp	
*	24059	29684: contig of 5626 bp	in length
*	29685	29784: gap of 100 bp	
*	29785	33951: contig of 4167 bp	in length
*	33952	34051: gap of 100 bp	
*	34052	39841: contig of 5790 bp	in length
*	39842	39941: gap of 100 bp	
*	39942	48297: contig of 8356 bp	in length

[illegible]

```
|||||
Db 16949 GGTCTTGAACCTCTGACCTCAGATGATCCGCCCTCGCCTCCCAAGTGCTGGATT 17008
Oy 1523 ACAGGcAtgagtcacccgcccgcgaaggtcagtgcttaataaggaataac 1575
Db 17009 ACAGGtGtAGCCACCgCGCCcAGcCTATTtTAATTtTTTtTtATCAAC 17061
```

Search completed: December 3, 2000, 18:25:29  
Job time: 43343 sec

